André Martinet

DES STEPPES AUX OCÉANS

L'indo-européen et les « Indo-Européens »



Payot

LANGAGES ET SOCIÉTÉS

Collection dirigée par Louis-Jean Calvet

De la langue aux langages la science se multiplie : non plus linguistique mais faisceau interdisciplinaire où les signes de nos sociétés s'éclairent.

La collection *Langages et sociétés*, à la croisée de ces démarches qui ont caractérisé les sciences humaines de notre époque, se veut un certain regard analytique : que pouvons-nous apprendre des sociétés qui nous entourent à travers les signes d'elles-mêmes qu'elles émettent ? Et de quelle utilité sociale, « militante », ce savoir est-il porteur ?

Déjà parus :

Marc ANGENOT : *La parole pamphlétaire.*
Sylvain AUROUX : *La sémiotique des encyclopédistes.*
Guy-Claude BALMIR : *Du chant au poème* (chant et poésie populaires des noirs américains).
Louis-Jean CALVET : *Les jeux de la société.*
— *Langue, corps, société.*
— *Chanson et société.*
André CHERVEL : *Histoire de la grammaire scolaire* (... Et il fallut apprendre à écrire à tous les petits Français) (P.B.P. 394).
Claude COUROUVE : *Vocabulaire de l'homosexualité masculine.*
Christian CUXAC : *Le langage des sourds.*
Yves DELAHAYE : *La frontière et le texte.* Pour une sémiotique des relations internationales.
— *L'Europe sous les mots* (le texte et la déchirure).
Gérard DELEDALLE : *Théorie et pratique du signe* (Introduction à la sémiotique de Charles S. Peirce).
Françoise ESCAL : *Espaces sociaux, espaces musicaux.*
Pierre GUIRAUD : *Dictionnaire érotique.*
— *Sémiologie de la sexualité.*
— *Dictionnaire des étymologies obscures.*
— *Structures étymologiques du lexique français.*
Nancy HUSTON : *Dire et interdire* (Eléments de jurologie).
MARX, ENGELS, LAFARGUE, STALINE : *Marxisme et linguistique* (précédé de L.-J. Calvet : *Sous les pavés de Staline la plage de Freud ?*).
Michel ORIANO : *Les travailleurs de la frontière* (bûcherons, cow-boys, cheminots américains au XIX[e] siècle).
Kyril RYJIK : *L'idiot chinois* (Initiation à la lecture des caractères chinois).
— *L'idiot chinois 2* (La promotion de Yu le Grand).
Maurice SWADESH : *Le langage et la vie humaine.*
J.-D. URBAIN : *La société de conservation* (Sémiologie des cimetières d'Occident).
Marina YAGUELLO : *Les mots et les femmes* (P.B.P. 397).

DES STEPPES AUX OCÉANS

DU MÊME AUTEUR

La gémination consonantique d'origine expressive dans les langues germaniques, Copenhague, Munksgaard, 1937.

La phonologie du mot en danois, Paris, Klincksieck, 1937.

La prononciation du français contemporain, Paris, Droz, 1945.

Economie des changements phonétiques, Berne, Francke, 1955.

La description phonologique, Genève, Droz, 1956.

Eléments de linguistique générale, Paris, A. Colin, 1960.

A functional view of language, Oxford, Clarendon, 1962, en français, *Langue et fonction*, Paris, Denoël, 1969.

La linguistique synchronique, Paris, P.U.F., 1965.

Le français sans fard, Paris, P.U.F., 1969.

Dictionnaire de la prononciation française dans son usage réel, avec Henriette Walter, Paris, France-Expansion, 1973.

Evolution des langues et reconstruction, Paris, P.U.F., 1975.

Studies in functional syntax, Munich, Fink, 1975.

Grammaire fonctionnelle du français, Paris, Didier, 1979.

Syntaxe générale, Paris, A. Colin, 1985.

ANDRÉ MARTINET

Professeur émérite à l'Université René Descartes
Directeur d'études à l'École pratique des Hautes Études

DES STEPPES AUX OCÉANS

L'INDO-EUROPÉEN ET LES « INDO-EUROPÉENS »



PAYOT, PARIS
106, Boulevard Saint-Germain
—
1986

A Henriette et Gérard Walter

PRÉFACE

pour les lecteurs

Le public cultivé, en France, ne s'est guère intéressé aux problèmes des origines indo-européennes. La découverte, il y a quelque deux cents ans, que la plupart des langues d'Europe, y compris le grec et le latin, dérivaient d'une même langue plus ancienne a pu remplir de satisfaction Allemands et Scandinaves dont les parlers natifs se voyaient, du coup, mis sur un pied d'égalité avec les langues classiques de l'Occident. Les Français, pour leur part, se sentaient les héritiers directs de ces langues et, de ce fait, peu concernés par cet élargissement préhistorique du concert des nations.

Les recherches archéologiques, en revanche, brillamment inaugurées en France, y ont toujours retenu l'attention. Mais comment mettre en rapport l'homme de Cro-Magnon, le Mas d'Azil et, plus récemment, les peintures de Lascaux avec des recherches de linguistique comparative, d'un haut degré d'abstraction, suscitées au départ par de lointains contacts avec des lettrés de l'Inde ? Pendant longtemps, les rapprochements esquissés entre les reconstructions linguistiques et les trouvailles faites dans des chantiers de fouille et des sites funéraires ont semblé, à certains même de ceux qui les risquaient, autant d'hypothèses gratuites. Aujourd'hui, l'extension, à grande échelle, des recherches archéologiques à la Russie et à l'ensemble des Balkans permet de restituer des processus d'expansion dans l'espace qui pourraient coïncider avec les déplacements de groupes d'usagers de langues indo-européennes. De telles identifications demeurent hypothétiques, mais elles ouvrent la porte à une confrontation cohérente des données livrées par les différentes disciplines. On peut désormais se risquer à présenter un cadre concret, avec localisations approximatives et chronologies relatives, où ordonner les faits de tous ordres.

Lorsqu'on se pose la question « l'indo-européen, où et quand ? » on

attend des réponses en termes de réalité géographique, historique et sociale. Mais on ne saurait oublier que l'indo-européen, qui délimite notre sujet, reste un concept purement linguistique. Si, dans notre sous-titre, les « Indo-Européens » apparaissent entre guillemets, c'est pour rappeler que les peuples qui, au cours des cinq derniers millénaires et jusqu'à ce jour, ont parlé ou parlent des langues indo-européennes ont peu de chances de représenter génétiquement ceux de qui, en dernière analyse, ils tiennent leur idiome. Les ancêtres linguistiques ne s'identifient pas nécessairement avec les ancêtres tout court. Les Français, qui se pensent volontiers comme des Gaulois auxquels on a imposé la langue latine, n'auront aucune difficulté à l'admettre.

Au sens large du terme, l'indo-européen est l'ensemble des langues passées, présentes et à venir qui dérivent, parallèlement, puis avec les divergences résultant de divers aléas, à partir d'un « tronc commun ». Fréquemment, on réserve le terme « indo-européen » à la désignation de ce tronc, même si l'on précise alors sa pensée par adjonction de l'épithète « commun » ou, en anglais, du préfixe *proto-*. Toutefois, cet indo-européen commun était, comme toutes les langues, dans un perpétuel état d'évolution, et, plutôt que la forme figée vers laquelle tendaient les efforts de nos prédécesseurs, c'est aujourd'hui un processus évolutif qu'on vise à restituer. Pour ce faire, on s'inspire non seulement du rapprochement des langues anciennement attestées, mais des antécédents que suggèrent les irrégularités de leur morphologie et des enseignements de la linguistique générale, structurale et fonctionnelle.

On voit mal comment on aurait pu, au cours de cet ouvrage, faire abstraction des données linguistiques. Affirmer simplement que le français et le russe ont été autrefois *la même langue* sera perçu comme une boutade, une façon désinvolte de jongler avec les faits. On saisira mieux ce qu'implique cette formulation si nous montrons que le mot *langue,* lui-même, avec ses deux valeurs d'organe de la déglutition et d'instrument de communication, dérive de *la même forme* ancienne que son équivalent russe *iazyk*, à l'issue de processus bien identifiés. Même si l'énumération de ces processus peut paraître austère, il a semblé qu'elle était nécessaire pour convaincre les lecteurs que l'identification de deux formes aussi différentes n'était pas un pur produit de l'imagination.

On attend donc, de beaucoup de ceux qui auront cet ouvrage entre les mains, une lecture « impressionniste » de certains passages. Le seul titre, « phonologie », du chapitre 9 ne manquera pas de faire peur à beaucoup. Celui de « grammaire », au chapitre suivant, n'est peut-être pas très alléchant. Mais, là encore, même si le lecteur ne s'astreint pas à tout enregistrer, un survol peut être informatif. Il convient, en la

matière, d'éviter les idées préconçues : tel s'étonnera de suivre aisément ce qui est dit des changements phonétiques, alors que maintes reconstructions grammaticales le laisseront rêveur.

On trouvera ci-dessous, aux pages 261-263, une table des conventions graphiques.

Il faut signaler en terminant que, sur bien des points, le tableau qui est présenté ici des données relatives à l'indo-européen s'écarte résolument de la tradition. Nous savons mieux aujourd'hui comment les langues fonctionnent et comment elles changent pour s'adapter aux besoins communicatifs de ceux qui parlent. Il faut donc aller de l'avant selon le principe qu'il vaut mieux s'exposer à l'erreur que de se condamner au silence par excès de précautions.

AVERTISSEMENT

aux chercheurs

Cet ouvrage reproduit, pour l'essentiel, l'enseignement que j'ai donné, à deux reprises, en 1976-1977 et 1983-1984, à l'Ecole pratique des hautes études. C'est à la demande de mes auditoires que je m'étais risqué hors du domaine de la linguistique générale impliqué par l'intitulé de ma direction d'études. Ces auditoires se recrutent essentiellement parmi des chercheurs qui joignent, à une formation de généraliste, une bonne pratique de la description des états de langue. En matière d'évolution linguistique, en l'absence d'une formation philologique et de la connaissance d'états successifs d'un même idiome, ils s'en tiennent le plus souvent à des notions théoriques. Beaucoup, cependant, désirent être informés des méthodes de la comparaison et de la reconstruction. Toutefois, dans leur ignorance de la plupart des langues anciennes mises alors à contribution, ils n'ont pas accès aux manuels classiques dont la lecture suppose une parfaite maîtrise du grec et du latin. C'est pourquoi, comme on pourra le constater, il est fait ici usage des langues modernes partout où elles permettent d'illustrer les phénomènes décrits : l'anglais *what*, en particulier dans sa prononciation américaine [hwɑt], rapproché du latin *quod*, témoigne, à lui seul, de trois des traits fondamentaux qui distinguent le germanique des autres langues de la famille.

Le second de mes cours a été intégralement enregistré et transcrit par mes amis Henriette et Gérard Walter, et l'essentiel de ce qu'on trouvera ci-dessous en est une version nettoyée des répétitions et des tournures familières qui marquent mon élocution, même universitaire. Au-delà de ce qui est généralement admis par les comparatistes contemporains, on retrouvera certaines des propositions que j'ai faites antérieurement et qui sont reproduites dans *Evolution des langues et reconstruction*, Paris, 1975. J'y ai développé en outre certaines suggestions que je n'avais présentées que sommairement dans quelque

compte rendu d'ouvrage, voire, simplement, dans une note ; ainsi l'interprétation des occlusives sonores comme le produit d'anciennes glottalisées.

Ce qui est neuf ici c'est, d'une part, la reconstruction d'un système de phonèmes prénasalisés qui permet d'expliquer, entre autres choses, les fameuses alternances de *-r* et de *-n-* et celles de *-m-* et de *-bh-* des cas obliques du pluriel, d'autre part, la présentation du processus de conjugaison qui a abouti à la constitution des différents systèmes verbaux des langues attestées et à l'identification progressive d'un ancien absolutif avec un ergatif en *-s* issu d'une construction ablative plus ancienne.

On relèvera sans doute un effort constant pour éviter la projection, sur l'écran d'un indo-européen commun, des produits de la comparaison, et pour tenter de reconstruire chaque fois un processus évolutif. Il ne m'échappe pas que l'entreprise est difficile et que son succès final réclamera qu'on enseigne aux jeunes linguistes à opérer avec la synchronie dynamique, celle qui permet l'observation des changements en cours.

Il me reste à remercier ceux qui ont bien voulu lire mon manuscrit et me faire profiter de leurs remarques critiques, à savoir, outre Henriette et Gérard Walter déjà mentionnés, Jeanne Martinet, Thérèse Martinet et Johann Öberg. Jean Haudry, en particulier, a droit à ma reconnaissance pour m'avoir fait profiter de sa compétence dans tous les domaines où je me suis aventuré. Il va sans dire que je reste seul responsable des hypothèses trop hardies ou des inconséquences qu'on pourra relever dans les lignes qui suivent.

CHAPITRE PREMIER

L'INDO-EUROPÉEN, OÙ ET QUAND

La plupart des mots dont nous nous servons quotidiennement sont susceptibles de prendre des sens différents selon les situations ou les contextes où on les emploie. Nous y sommes habitués et cela ne nous gêne pas. La situation est différente lorsqu'on aborde l'étude attentive d'un phénomène. Il est alors indispensable de caractériser exactement ce dont on va traiter et si l'on se sert pour le désigner d'un terme de la langue courante, il faudra préciser lequel de ses divers sens est retenu en l'occurrence. Dans bien des cas, il s'impose de trouver, pour l'objet d'étude, un terme nouveau et c'est bien ce qui s'est produit pour l'indo-européen. Ceux qui ont les premiers utilisé ce terme savaient sans doute quelle valeur ils lui attribuaient. Mais, comme il s'agissait au départ d'une vue de l'esprit, il n'est pas étonnant que le terme ait pu, au cours du temps, se charger d'implications diverses qui varient selon les époques, les chercheurs et les auteurs. On ne devra donc pas s'étonner si nous tentons en priorité de définir les termes que nous avons retenus dans notre sous-titre.

Il faut, dès l'abord, distinguer entre l'adjectif *indo-européen* et le substantif correspondant. L'adjectif *indo-européen* a été créé pour caractériser des langues que l'on estime issues, par évolution régulière, d'une même langue disparue et non attestée. Ces langues, longtemps limitées à l'ancien continent, de l'Atlantique au Golfe du Bengale, sont aujourd'hui parlées, comme langues premières ou secondes, dans le monde entier.

Il y a donc de nombreuses langues indo-européennes parmi lesquelles le français, l'anglais, l'allemand, le russe et d'autres moins connues, comme le bengali ou l'ossète.

Le substantif *indo-européen*, créé à partir de cet adjectif, désigne, le plus souvent, la langue non attestée elle-même. On peut entendre des énoncés comme « En indo-européen, « cheval » se disait *ekwos* ». Mais

les comparatistes sont, dans ce cas, plus précis et parlent d'indo-européen commun, en anglais de *Proto-Indo-European*, en allemand, de *Urindogermanisch*, les Allemands préférant employer *Indogermanisch*, plutôt que son équivalent international *Indoeuropäisch*. Cet indo-européen avait, au départ, été conçu comme peu différent du sanskrit, langue sacrée des brahmanes et langue de culture de l'Inde. C'est la découverte du sanskrit par les Européens, au XVIII[e] siècle, qui avait suscité son rapprochement avec les langues classiques de l'Occident et donné naissance à la linguistique comparée. La pratique de cette nouvelle discipline a vite fait comprendre qu'il ne pouvait y avoir identité entre le sanskrit et la langue commune dont on postulait l'existence dans un lointain passé. On a de ce fait été tenté de reconstruire cette dernière en retenant, parmi les traits attestés dans les différentes langues en cause, ceux qui étaient les plus répandus et les moins susceptibles de correspondre à des innovations particulières. C'est ainsi que pour « père », on restituait une forme *pətēr*, avec un *p* initial majoritaire, celui du latin *pater*, du grec *patēr*, du sanskrit *pitā(r)* et l'on interprétait le *f* du germanique (anglais *father*), le *h* de l'arménien (*hayr*) et le zéro du celtique (irlandais *athir*) comme des affaiblissements ultérieurs. Dans un second temps, par un effort d'interprétation des structures reconstruites, on est remonté au-delà de l'équation *pətēr* = « père ». D'abord, sur le plan de la forme, en interprétant le -ə- de la forme reconstruite comme une réduction, en syllabe inaccentuée, d'un groupe primitif *-eH-* où *H* est, en quelque sorte, une grandeur algébrique. Ensuite, sur le plan du sens, en posant, pour *pətēr*, non plus la valeur de « géniteur », mais celle de chef de la grande famille patriarcale.

La découverte, au début du XX[e] siècle, de nouvelles langues, comme le tokharien et le hittite, évidemment apparentées à celles qu'on considérait comme indo-européennes, est venue renforcer la conception d'un indo-européen évoluant dans le temps, comme toutes les langues, en suggérant la possibilité, voire la vraisemblance, de séparations successives à partir du tronc commun. Dans ces conditions, la reconstruction ne peut plus être celle d'un état de langue conçu comme parfaitement stable, à l'image de ce que sont souvent, dans notre imaginaire, nos langues nationales contemporaines. Elle devient celle d'un processus évolutif qui se poursuit jusqu'à nos jours depuis la nuit des temps. Cela amène à remettre en cause la notion d'un « indo-européen commun » qu'on pourrait, avec le secours des archéologues, localiser, plus ou moins précisément, dans le temps et dans l'espace. Il nous faut désormais tenter de l'imaginer sous la forme d'une langue en perpétuelle évolution, celle d'une communauté constamment exposée à voir certaines de ses composantes faire sécession et s'établir ailleurs. Parmi les autres, celles qui ne bougent

pas, peuvent s'établir, dans une région particulière, des contacts privilégiés qui, en matière de langue, vont entraîner des innovations particulières. Il en résultera, d'une part, l'apparition de dialectes différents, mais concurremment, l'élimination de divergences qui avaient commencé à se fixer. Ces flux et ces reflux sont de toutes les époques et l'on ne peut imaginer, sur l'axe du temps, un point avant lequel existait une langue immuable et homogène et après lequel se trouvaient séparées et parfaitement distinctes les langues attestées plus tard par des textes ou encore parlées jusqu'à ce jour. Devons-nous, en conséquence, renoncer à construire des formes disparues ? En aucune façon. Mais chaque fois que nous le faisons, nous devons rester conscients que chacune d'entre elles est, au mieux, une étape : la forme *pətēr*, que nous avons présentée ci-dessus, n'est pas à écarter, mais à poser comme un intermédiaire entre une plus ancienne notée *pH°ters* et celles que nous trouvons plus tard dans les textes anciens et dans nos langues contemporaines. Si l'on hésite à prononcer *pətēr* et, plus encore, *pH°ters*, c'est que ces graphies ne prétendent pas reproduire une réalité orale précise. Elles ne constituent chacune qu'une sorte de formule à partir de laquelle le comparatiste peut retrouver les diverses formes attestées et les relations qu'on postule entre ce terme et d'autres termes reconstruits de la même façon.

Il faut bien comprendre que le passage d'une étape à la suivante n'implique pas nécessairement que le domaine des langues indo-européennes va se trouver fractionné : si nous mettons à part un indo-européen occidental, c'est simplement que nous sommes mal informés quant aux formes linguistiques pratiquées par des populations situées géographiquement entre ces Occidentaux et les ancêtres linguistiques du grec ou du slave. En d'autres termes, cette esquisse chronologique reflète plus l'état de nos connaissances que la réalité des faits.

Lorsque s'élargira le domaine indo-européen, il s'y produira nécessairement des solutions de continuité, c'est-à-dire de vastes espaces où vont continuer à se parler d'autres langues. On peut envisager que, pendant longtemps, l'indo-européen va rester, dans tel ou tel coin du monde, la langue d'une classe dominante vivant au contact d'une majorité d'allophones. Mais cela n'entraîne pas nécessairement que les contacts soient rompus entre les classes dominantes des différents îlots ainsi formés. Lorsque, dans la première moitié du premier millénaire avant notre ère, les Celtes restent encore, dans ce qui sera plus tard la Gaule et l'Espagne, des nouveaux venus probablement minoritaires, les liaisons se maintiendront, à travers les Pyrénées et le pays des Aquitains, cousins des Basques d'aujourd'hui, par une route ponctuée de forteresses, les *Verdun* de la toponymie celtique.

Lorsqu'on désire, légitimement, préciser les réponses aux

« quand » en datant les hypothèses, et tenter de se prononcer sur les lieux où l'on doit situer telle ou telle variété de l'indo-européen à telle date, il faut nécessairement avoir recours, d'une part, à une reconstruction sémantique où l'on va chercher à se représenter les choses et les notions qui correspondent aux unités linguistiques, d'autre part aux résultats de la recherche archéologique.

Beaucoup de comparatistes sont longtemps restés sceptiques quant à la possibilité de localiser les reconstructions linguistiques avec le secours de données extérieures à leur discipline. Les réticences en la matière se comprennent fort bien : on répugne naturellement à se risquer hors du domaine de sa compétence. Les philologues, que sont fréquemment, au départ, les comparatistes, sont tentés, lorsqu'ils cherchent à interpréter les faits linguistiques, de partir de ce qu'ils connaissent de la culture des civilisations classiques, plutôt que d'utiliser les données de l'anthropologie et de l'archéologie contemporaines.

De leur côté, les données archéologiques doivent être interprétées en termes évolutifs. Chaque culture est à envisager comme un procès en voie d'expansion ou de récession, et un effort doit être fait pour rétablir, là où elle existe, la liaison entre les mouvements préhistoriques, les développements bien attestés et les processus observables dans le monde contemporain. La conquête du monde, par les peuples de langues indo-européennes qui ont su jusqu'ici garder une avance dans la mise de supériorités techniques au service de la violence, a commencé par la subjugation des populations préexistantes, de l'Inde à l'Irlande. Elle ne s'est pas terminée avec la conquête de l'Ouest américain et de l'Asie septentrionale et l'impérialisme colonial, car elle se prolonge de nos jours avec la terreur atomique. Il y a eu des péripéties diverses, des avancées et des reculs, mais aucune réelle solution de continuité depuis les expéditions printanières lancées à la conquête de l'Italie au cours du millénaire avant notre ère, jusqu'aux cavaliers de Hernán Cortés, aux divisions blindées des années quarante de ce siècle et au napalm du Vietnam.

On ne peut commencer à comprendre le phénomène indo-européen que si l'on cesse de l'aborder sous les seules espèces des grandes cultures du passé dont chacune représente déjà un amalgame. Il faut tenter de dépasser la conception, combien tenace, d'une seule et unique diaspora indo-européenne avec, d'emblée, un nombre déterminé de nouveaux peuples correspondant à ceux que nous connaissons bien parce qu'ils ont laissé des traces dans l'histoire.

A côté de peuples, comme les Vénètes d'Italie et les Messapiens, dont il nous reste, outre le nom, quelques inscriptions qui nous convainquent qu'ils étaient bien de langue indo-européenne, combien d'autres, de langues apparentées, ont dû disparaître dont nous ne

savons rien, subjugués ou absorbés par leurs voisins ! Combien de nouveaux contacts ont pu s'établir qui ont rapproché des parlers au départ fort divergents au point d'en faire des variétés d'une même langue ! Nous verrons plus loin quelle place il faut accorder à la convergence des langues pour expliquer l'assimilation de migrants à leur nouveau milieu.

Avant d'essayer de répondre aux deux questions qu'on pose, dans le titre de ce chapitre, il faut en formuler une troisième : qui a parlé ou parle l'indo-européen ?

A ne considérer que les indices proprement linguistiques, c'est-à-dire en s'en tenant aux parties nettement structurées des langues et sans faire intervenir les signifiés qui gagnent à être rapprochés des données archéologiques, on peut tenter de répondre à « qui », « où » et « quand » si l'on veut bien se contenter d'une chronologie relative à peine datée.

Dans un premier temps, dans la préhistoire, le terme d'indo-européen s'applique à la langue des ancêtres (linguistiques) de *tous* ceux que nous identifions aujourd'hui comme ayant appartenu ou appartenant au groupe. Il faut naturellement préciser « ancêtres linguistiques », parce que, aux dates historiques, y compris la nôtre, il y a toute chance pour que ceux qui parlaient, ou parlent, indo-européen aient eu une majorité d'ancêtres qui parlaient d'autres langues.

Dans un second temps, pour lequel on dispose d'une documentation linguistique, on distinguera, tout d'abord, entre 1) les *Anatoliens*, ceux qui sont établis en Asie Mineure aux troisième et deuxième millénaires avant notre ère mais qui ont pu faire sécession beaucoup plus tôt, représentés essentiellement par les Hittites, et 2) tous les autres, ceux qui restent assez étroitement groupés pour innover ensemble, par exemple, en établissant une distinction entre un genre féminin et un genre masculin.

Dans un troisième temps, s'esquisse, parmi ces derniers, une distinction entre 1) les populations qui, à l'est, vont palataliser leurs dorsales, c'est-à-dire, par exemple, changer le *g-* qu'on trouve dans le latin *gnosco* « connaître » en un *z* ou en l'équivalent du français *j* dans *joli* ou *dj* dans *adjoint*, et 2) celles qui, au moins pour l'instant, conservent intacts leur *k* et leur *g* et qui sont localisées en général plus à l'ouest. Les premières comprendront les Indo-Iraniens qu'on retrouvera plus tard en Asie, de la Mésopotamie au Golfe du Bengale, les Arméniens, les Albanais, les Baltes et les Slaves. Le grec, qui a échappé à la palatalisation, où donc *(gi)gnosko* « connaître » a gardé son *g*, a dû cependant rester longtemps en étroit contact avec les parlers de ceux qui vont devenir les Indo-Iraniens.

L'indo-européen occidental, qui s'étendra finalement de la Baltique

à la Méditerranée et à l'Atlantique, porte en germe les variétés qui deviendront l'italique, le celtique et le germanique.

Pendant trop longtemps, les efforts pour répondre à la question « où » ont été irrémédiablement faussés par des préjugés nationalistes, chacun désirant retrouver un habitat indo-européen primitif le moins loin possible de sa propre résidence : on ne reviendra pas ici sur l'insistance de certains chercheurs allemands pour mettre en rapport un tel habitat avec, par exemple, la présence du hêtre, ce qui tendait à repousser à l'ouest le domaine originel des « Indo-Européens ». Quant aux Français, qui avaient dû, dès l'abord, renoncer à retrouver les origines indo-européennes entre la Meuse et les Pyrénées, ils se sont généralement désintéressés de ce problème.

Les progrès réalisés au cours des dernières décennies par les sciences archéologiques, notamment en matière de datation du produit des fouilles, et l'exploration rationnelle de nouveaux sites, permettent aujourd'hui, pour la première fois dans l'histoire de la recherche, d'établir, avec quelque chance de retenir l'attention, sinon nécessairement de convaincre, une liaison entre données linguistiques et données archéologiques.

Le premier effort dans ce sens est lié au nom de Marija Gimbutas et c'est essentiellement de ses écrits[1] que je m'inspire dans l'esquisse qui va suivre. Je n'ai pas ici l'intention de justifier chacune de mes assertions. Je ne m'adresse pas, pour l'instant, aux spécialistes de la reconstruction qui seuls seraient habilités à discuter les points de vue adoptés. Toute théorie scientifique est faite pour susciter des mises au point et des théories adverses. Mais sa présentation à un public qui ne se limite pas aux spécialistes ne peut viser qu'à offrir une base de réflexion et un cadre dans lequel pourront s'ordonner certaines données. Le schématisme, voire le manichéisme, est ici volontaire. Chacun de mes lecteurs saura nuancer à sa guise les images simplistes et parfois brutales que j'apporte.

5000 ans avant notre ère, le peuple de langue indo-européenne est localisé au sud-est de la Russie d'aujourd'hui, dans la région dite des kourganes. Les kourganes sont des tumulus où l'on retrouve les restes

1. Voir essentiellement Die Indoeuropäer, dans *Die Urheimat der Indogermanen*, Darmstadt, dir. A. Scherer, 1968, p. 538-571, deux articles dans *The Journal of Indo-European Studies*, vol. 5, 1977, p. 277-338, The first wave of Eurasian Steppe Pastoralists into Copper Age Europe, et vol. 8, 1980, p. 273-315, The Kurgan wave 2. (*c.* 3400-3200 B.C.) into Europe and the following transformation of culture, et finalement Old Europe in the fifth millennium B.C., The European situation on the arrival of Indo-Europeans, dans *The Indo-Europeans in the fourth and third millennia*, sous la direction d'Edgar C. Polomé, Ann Arbor, Michigan, Karoma, 1982, p. 1-60. On trouvera dans ce volume, cité désormais sous l'abréviation *The Indo-Europeans*... des suppléments bibliographiques. Il est fait état de tout cela dans l'article Les chemins de l'orgueil, de Patrice Leclercq paru dans *le Monde* du 28.1.83.

de ce qu'on suppose être un chef, entouré de richesses, souvent somptueuses, et des squelettes d'un certain nombre de jeunes femmes et de serviteurs. Si l'on oublie un instant les théories visant à rendre compte, en termes sociologiques ou religieux, de la mise à mort des veuves et des proches du défunt, on comprend l'efficacité de ces pratiques pour prévenir toute tentative criminelle de la part de l'entourage. Nous avons là affaire à une société patriarcale hautement hiérarchisée, celle que laisse attendre la reconstruction sémantique. Des sépultures du même type vont se retrouver à travers l'Europe, jusque dans ce qui est aujourd'hui l'Allemagne moyenne. Mais, au fur et à mesure qu'on se déplace vers l'ouest, les datations sont plus récentes et l'on trouve moins de richesses et moins de personnes immolées. Cela implique une poussée conquérante dans la direction de l'occident à travers des régions où les sépultures antérieures témoignent généralement de conceptions plus égalitaires, dans la mort, sinon dans la vie, où chacun a sa propre tombe et où l'on meurt seul.

Le patriarcat des kourganes laisse supposer un panthéon essentiellement masculin reflétant les trois états de la société : les prêtres, les guerriers et les pasteurs. Il s'agit en effet, au départ, de nomades plus enclins à mettre à profit ce qu'ils rencontrent dans leurs errances qu'à susciter eux-mêmes, par la culture, l'apparition de nouvelles ressources. Ce sont des éleveurs, certes, mais ils restent, dans un sens, des prédateurs. Le cheval, plus attelé que monté, jouera un rôle important dans leur expansion.

Vers l'ouest, en trois vagues successives qu'on date d'avant – 4000 à après – 3000, donc sur un millénaire et demi environ, le peuple des kourganes va pénétrer dans ce que nous connaissons aujourd'hui comme la plaine du Danube et les Balkans. Il y trouvera une civilisation avancée, de type matriarcal et agricole, donc culturellement aussi différente qu'il est possible de la leur, avec un culte et des déesses de la fécondité. Certes, les Indo-Européens vont finalement imposer leur langue et certains hôtes de leur panthéon, mais non sans que soit produit l'amalgame qui nous vaut, en face des dieux du tonnerre et de la guerre, des déesses comme Gê, Déméter, Perséphone, Athéna, à côté des Vénus, des Junon et des Freyja de sociétés où la femme est conçue, en priorité, comme « le repos du guerrier » ou la protectrice des héros.

Il reste à voir à quel échelon de cette préhistoire on doit placer la séparation des Anatoliens, puis celle, plus tardive, des Indo-Iraniens.

On peut se demander par quelle route ils ont, à partir des steppes des kourganes, atteint l'Asie mineure. Pour les Anatoliens, qu'on y trouve établis à demeure, on pourrait penser à un cheminement par le Caucase et les bords orientaux de la Mer Noire. Pour les Indo-Iraniens, pour qui l'Asie mineure pourrait n'être qu'une étape dans

leur migration vers l'Iran et l'Inde, on verrait plutôt un passage à travers les Balkans, ce qui s'accorderait bien avec les analogies structurales qu'on relève entre grec et sanskrit. Quoi qu'il en soit, les hypothèses qu'on présente, en la matière[2], paraissent plus sujettes à caution que celles qui concernent l'indo-européanisation de l'Europe.

Que penser, maintenant, des Indo-Européens avant – 5000 ? Nicolas Troubetzkoy[3] a présenté l'hypothèse qu'ils étaient le produit de l'amalgame de différentes populations. Cela permettrait peut-être d'expliquer certaines hétérogénéités linguistiques. Il est frappant, par exemple, que, dans la numération de 1 à 100, la seule qu'on puisse reconstruire pour un indo-européen ancien — Anatoliens mis à part, pour lesquels la documentation nous manque —, on ne rencontre, comme consonnes occlusives, que celles des séries désignées traditionnellement comme sourdes et sonores, à l'exclusion des « sonores aspirées ». Le grec attique, par exemple, présente, de 1 à 10 *heîs, dúo, treîs, téttares, pénte, héx, heptá, oktṓ, ennéa, déka*, pour 20, *eíkosi*, pour 100, *hekatón*. Dans cette liste, il n'y a pas trace des *ph*, *th* et *kh* qui attesteraient la présence des occlusives aspirées. Or, dans l'ensemble du vocabulaire, ces aspirées ont une fréquence analogue à celle des sourdes et nettement supérieure à celles des sonores simples. Cela laisse supposer que la numération provient d'une langue autre que celle qui a fourni l'essentiel de l'effectif du lexique, mais suggère plutôt l'emprunt de ce système que l'amalgame de deux populations de langues différentes.

Vittore Pisani[4] envisagerait, au départ, la rencontre de nomades guerriers errant dans les steppes et de prêtres d'origine caucasique. La tentation, fréquente, de faire intervenir le Caucase pourrait s'expliquer simplement par la proximité géographique. Mais, en fait, elle dérive surtout du désir de justifier certaines hypothèses relatives aux structures phonologique et syntaxique que l'on postule pour des stades très anciens de la langue.

Comme on le verra par la suite, on pense aujourd'hui que les consonnes sonores, du type de *b*, *d* et *g*, qu'on rencontre, par exemple, en français, proviennent de glottalisées[5], c'est-à-dire d'articulations accompagnées d'une fermeture de la glotte. Or, les langues caucasiques sont, dans les régions occidentales de l'Ancien Continent, les seules qui présentent aujourd'hui les glottalisées. Toutefois, les

2. Voir, par exemple, Homer L. THOMAS, Archaeological evidence for the migrations of the Indo-Europeans, *The Indo-Europeans*..., p. 61-86.

3. Dans Gedanken über das Indogermanenproblem, *Acta linguistica* I, p. 81-89.

4. Dans *Indogermanisch und Europäer*, Munich, 1974.

5. Cf., dans André MARTINET, La palatalisation « spontanée » de *g* en arabe, B.S.L., 54 (1959), p. 90-102, reproduit dans *Evolution des langues et reconstruction*, Paris, P.U.F., 1975, p. 233-247, la description du passage de *k* glottalisé à *g*, p. 241.

recherches contemporaines ont démontré la fréquence de ce type dans les systèmes à trois séries d'occlusives, tel celui qu'on pose aujourd'hui pour l'indo-européen ancien. En conséquence, la nécessité de rattacher l'indo-européen à une aire de glottalisation englobant le Caucase ne s'impose nullement. Rappelons qu'on pose également des glottalisées à la source des « emphatiques » des langues sémitiques contemporaines[6].

En matière syntaxique, on a depuis longtemps envisagé, pour les stades très anciens de l'indo-européen, une construction ergative, c'est-à-dire l'emploi d'une marque positive pour désigner l'agent lorsqu'il figure dans le même contexte qu'un patient, alors que ce dernier apparaît sans aucune indication de fonction, ni désinence ni particule. Sur ce point également les langues caucasiques présentent un trait postulé pour un stade archaïque de l'indo-européen. Mais la construction ergative est un phénomène très répandu à travers le monde et qui, à la réflexion, paraît aussi « logique » que la construction avec objet à laquelle nos langues européennes contemporaines nous ont habitués[7].

Nous n'excluons nullement la possibilité de contacts entre les anciens Indo-Européens et leurs contemporains caucasiques, mais on n'en a jusqu'ici relevé aucune trace précise. Les similitudes structurales qu'on pourrait envisager à une date très reculée n'impliqueraient ni communauté d'origine ni période de symbiose.

Un apparentement soit avec les langues chamito-sémitiques pratiquées aujourd'hui du Golfe Persique au Maghreb, soit avec les langues finno-ougriennes parlées en Hongrie, dans l'Eurasie septentrionale, voire, comme on cherche à le démontrer[8], par les Indiens de la Californie centrale, a, depuis longtemps, été envisagé. Mais tout ceci reste dans le domaine des hypothèses invérifiées et, peut-être, invérifiables[9].

Nous avons, dans les pages qui précèdent, rencontré des formes reconstruites par comparaison et par hypothèse. On en rencontrera fréquemment dans ce qui va suivre. Elles y seront, selon la tradition, précédées d'un astérisque qui signale que ces formes ne sont attestées dans aucun des textes existants, manuscrits ou inscriptions lapidaires.

6. Voir, par l'auteur de ces lignes, l'exposé intitulé Remarques sur le consonantisme sémitique, dans B.S.L. 49 (1953), p. 67-78, reproduit dans *Evolution...*, p. 248-261.

7. Voir, par ex., André MARTINET, *Syntaxe générale*, Paris, 1985, § 8-14 à 8-18.

8. Des recherches, menées par Otto J. SADOVSZKY, débutent en 1976 avec Report on the state of Uralo-Penutian Research, *Ural-Altaische Jahrbücher*, Wiesbaden, 48, p. 191-204. A paraître Contributions to an Ob-Ugrian-Maiduan Comparative Grammar.

9. A signaler l'audacieuse tentative de N. D. ANDREEV, dans *Ranneindoevropejskij prajazyk*, Leningrad, 1986, pour faire dériver l'indo-européen d'un proto-boréal qui serait à l'origine des langues de l'Eurasie septentrionale.

Si l'on tente de les prononcer, on se rappellera que les lettres y ont le plus souvent la même valeur qu'en français. Toutefois *u* se prononce comme le *ou* de *chou* et le *e* comme *é* dans *été* ou *è* dans *après*. Dans *ei*, *eu*, *oi*, *ou*, les deux lettres se prononcent. Le signe *ə* peut être reproduit comme le « *e* muet » de *je n'veux pas*. Un petit *o* au-dessus de la ligne, qui vaut, en général, pour « degré », correspond à une voyelle analogue, aussi brève que possible. On attribuera aux *H* des prononciations différentes selon le chiffre placé en indice : H_2 pourrait être reproduit comme le *ch* de l'allemand *Bach* ou le *j* de l'espagnol *jamás* ; H_3 comme le groupe *ju* de l'espagnol *Juan*. Quant à H_1, on pourra, en pratique, le traiter comme un *h* muet.

Les notations phonétiques sont faites au moyen de l'alphabet phonétique international. Pour les articulations chuintantes on a donné la préférence aux formes surmontées du circonflexe renversé : [š], [ž], correspondant respectivement au *ch* et *j* du français. Pour la consonne initiale de l'anglais *thin* ou de l'espagnol *cinco*, on a préféré le *thorn* germanique Þ, qui va, en tout état de cause, apparaître dans des formes citées, au grec *thêta* que les élèves des classes désignent souvent comme « le cigare ». On s'est permis quelques entorses au système lorsqu'on a pensé que les lecteurs, habitués aux valeurs françaises des lettres, s'y retrouveraient mieux : pour l'équivalent russe de « langue », [iazyk] a paru préférable à [jazyk]. Les notations placées entre crochets carrés reproduisent le détail des prononciations ; entre barres obliques, elles reproduisent les unités distinctives de la langue. Sont normalement données en *italique* aussi bien les formes à astérisque que celles empruntées aux diverses langues citées. Celles qui proviennent de langues écrites au moyen de l'alphabet latin sont reproduites telles quelles. Celles qui sont normalement écrites dans un autre alphabet sont translittérées, avec une lettre (ou parfois deux) de l'alphabet latin remplaçant une lettre de l'autre alphabet. Pour les langues à écriture idéographique, comme le japonais ou le chinois, il existe des équivalents officiels en caractères latins.

Pour plus de précisions relatives aux types articulatoires et aux notations phonétiques, on se reportera au tableau des pages 261-263.

Lorsqu'on cite des formes indo-européennes, reconstruites ou anciennes, on le fait souvent parce que l'on traite du radical du mot, sans s'occuper des finales qui varient selon la fonction de ce mot dans la phrase. Pour marquer qu'il s'agit dans ce cas du radical, on le reproduira avec un trait d'union final : **newo-* « neuf, nouveau ». **Newo*, sans trait d'union, renvoie à une époque où le mot pouvait apparaître nu, sans finale. On fera souvent usage du trait d'union pour visualiser l'analyse d'un mot en ses éléments composants : **owi-o-m* composé de *owi-* « mouton », du suffixe adjectival *-o-* et de la finale *-m* de nominatif-accusatif neutre.

L'adverbe « régulièrement » utilisé en référence à un changement phonétique veut dire que, dans la langue et à l'époque en cause, il était inévitable, donc prévisible. C'est son absence qui réclamerait qu'on recherche un conditionnement particulier, comme, par exemple, l'analogie d'une autre forme. Nous ne pouvons, bien entendu, pas offrir, pour chacun de ces changements la documentation qui en confirmerait le caractère inéluctable et nous prions simplement nos lecteurs de nous faire confiance. Nous reviendrons, ci-dessous, au chapitre VII, sur ces problèmes.

Il n'est pas facile de dater exactement les différents événements qui ponctuent l'expansion indo-européenne et les innovations langagières qui ont abouti à faire, d'une seule et même langue primitive, une foule d'idiomes distincts. Lorsqu'on se réfère à des siècles ou à des millénaires, on s'en tire en général en utilisant des numéros d'ordre chronologiques : le XVe siècle, le deuxième millénaire, par exemple, en marquant, chaque fois que cela ne ressort pas du contexte, s'il s'agit de notre ère ou des âges qui l'ont précédée. Lorsqu'on peut spécifier une année, comme datation soit précise, soit approximative, on fera précéder du signe moins les années avant notre ère : – 106 par exemple, veut dire 106 années avant la date présumée de la naissance de Jésus-Christ, c'est-à-dire six ans avant la fin du deuxième siècle compté à rebours à partir de cette date.

CHAPITRE II

SUBSISTANCE ET DÉPLACEMENTS DE POPULATION

L'impossibilité où nous nous trouvons de choisir, sur l'axe du temps, un point pour y reconstruire un état de langue, une communauté qui en ferait usage et une culture correspondante nous impose une vision dynamique du phénomène indo-européen. Plus donc que les périodes de stabilité relative, ce qui va retenir notre attention sont les processus évolutifs, qu'ils soient linguistiques ou culturels, et, bien entendu, les déplacements qui vont rompre les liens existants et établir de nouveaux contacts entraînant divergences d'un côté, convergences de l'autre.

Selon la vision historique traditionnelle, les mouvements de population sont conçus, sinon nécessairement comme des catastrophes, au moins comme des épisodes regrettables qui viennent affecter la tranquillité de sociétés bien assises. Même lorsqu'ils résultent d'une initiative des gens en place, comme la venue, en Europe occidentale et centrale, d'allogènes maghrébins, turcs ou autres, au cours du troisième quart du XX[e] siècle, ils créent tôt ou tard des problèmes. On pense aux Germains, établis pacifiquement comme colons dans l'Empire romain, qui se sont révélés comme la première amorce d'un déferlement à venir. L'histoire est, le plus souvent, celle du pays qui a pâti de l'arrivée en masse de nouveaux venus et l'on comprend qu'elle traite plus volontiers des conséquences de cette arrivée que des circonstances qui ont déterminé le mouvement. Notre point de vue est nécessairement tout autre puisque notre attention se centre, au départ, sur les envahisseurs, même si l'identité, la culture et le sort des populations en place sont loin de nous être indifférents. En d'autres termes, nous devons en priorité nous demander pourquoi les peuples de langue indo-européenne se sont mis en marche.

Si l'on rencontre tant de chercheurs qui répugnent à rechercher le pourquoi des choses c'est que tout phénomène, quelque insignifiant

qu'il soit, se révèle comme le point d'incidence d'une infinité de causes dont la plupart échappent vite aux possibilités de vérification, voire simplement d'examen. Il y a pourtant un moyen d'aborder avec fruit l'étude des causes. C'est d'en identifier une probable et de chercher, dans chaque cas, dans quelle mesure elle a pu intervenir pour susciter ou simplement favoriser le phénomène.

Soit, par exemple, l'expansion islamique qui a suivi l'Hégire, aux VIIe et VIIIe siècles de notre ère, et amené les Arabes jusqu'à l'Atlantique. A première vue, elle paraît déterminée par l'éclosion d'une foi nouvelle, source d'un intense prosélytisme. Mais, à moins que nous acceptions d'en trouver la cause première et unique dans la volonté d'Allah, nous devons supposer un complexe de conditions particulières, la prédication de Mahomet n'en étant qu'une, qui a donné une extraordinaire ampleur à l'un des multiples débordements humains qu'a produits l'Arabie. Cette péninsule, à moitié désertique, mériterait, autant ou plus que la Scandinavie, la désignation d'*officina gentium* [1], d'usine à fabriquer des peuples. Elle est le type même de la vaste région, à ressources naturelles limitées, située à proximité d'un pays riche, ici la Mésopotamie, creuset de successives civilisations, d'où rayonnent les innovations susceptibles d'atténuer, puis de réduire les famines. Les disettes récurrentes étaient précisément le contrepoids à la fécondité naturelle, celle qui fait naître des enfants aussi longtemps que la femme en peut porter. Leur atténuation, voire leur disparition, aboutit à ce qu'on désigne comme un boom démographique, une croissance d'une extraordinaire rapidité, prélude possible d'une expansion.

C'est une situation de ce type qu'on peut supposer en Europe, vers le début de notre ère, avec un Empire romain concentrant, puis diffusant une masse d'innovations techniques qui aboutissent à une mise en valeur plus satisfaisante des ressources naturelles et, en conséquence, à une amélioration générale des conditions de vie. Il en résulte, chez les voisins, Basques, Germains et autres, qui ne sont pas affectés par la dénatalité des populations romanisées et qui connaissent la fécondité naturelle, une expansion démographique considérable qui exerce sur les frontières une pression constante. Dès que se retirent les légions romaines, c'est le débordement qu'on a désigné comme les invasions barbares. A plus petite échelle, les Basques non-assimilés des zones escarpées et boisées des Pyrénées et de Cantabrie déferlent jusqu'à la Garonne pour former la Gascogne (*Vasconia*) et, vers le sud-ouest, jusqu'aux abords de Burgos où ils vont contribuer de façon

1. L'expression est de Jordanes, évêque et historien des Goths au VIe siècle, à qui nous devons l'essentiel de ce que nous savons sur l'histoire de ce peuple.

décisive à faire du castillan une langue profondément distincte des autres parlers de la Péninsule[2].

Dans le monde contemporain, les progrès de la médecine relaient ceux des techniques rurales pour faire reculer les disettes et maintenir en vie la foule des enfants qu'on rencontre partout où ne s'est pas établi le contrôle des naissances. D'où les démographies galopantes qu'on constate dans le Tiers-Monde. Les famines, nous le savons, ne sont pas exclues là où interviennent des circonstances climatiques particulièrement défavorables : les humains meurent par milliers au Sahel. Toutefois, l'amélioration générale des conditions de subsistance peut maintenir en vie une population assez dense dans les pays dits « en voie de développement ». Cependant la différence entre la richesse des pays de tête et le dénuement relatif de ceux qui suivent crée un déséquilibre que les restrictions à la libre circulation ne permettent pas de résorber. Ce déséquilibre peut être à la source de violences diverses.

En résumé, l' « invasion barbare » ou, en d'autres termes, l'irruption de conquérants dans un pays de niveau de vie supérieur, a toute chance de résulter d'une explosion démographique chez une population qui a profité d'innovations techniques en matière de production de biens de consommation... et d'outils d'agression, sans être affectée par la limitation des naissances. Mais, dans les pays qui sont en tête du développement, il n'est pas acquis que s'étendent à l'ensemble de la population la connaissance et la pratique des moyens anticonceptionnels. Il pourra donc se produire des excédents de population qu'il sera tentant d'exporter. D'où les colonies de peuplement avec élimination physique des autochtones, leur réduction à l'état de serfs ou leur refoulement dans les zones les moins fertiles du territoire. Les colonies d'exploitation et la conquête de nouveaux marchés complètent le panorama de l'expansion dans le monde contemporain. Elles ont sans doute joué un moindre rôle aux époques dont nous traitons ici.

Les conséquences linguistiques d'un mouvement de population dépendent, bien entendu, de facteurs divers parmi lesquels il faut retenir, en premier lieu, la puissance respective des groupes humains qui vont se trouver en contact : des gens peu nombreux et démunis n'ont aucune chance d'imposer leur langue lorsqu'ils se trouvent en contact avec une population abondante et aisée. On peut être tenté de faire valoir le degré de culture comme un des facteurs qui vont contribuer à faire triompher une langue sur une compétitrice : le roman a pu, en Gaule, l'emporter sur le germanique parce qu'il était le vecteur de la culture latine ; mais, pour expliquer les frontières des

2. Voir André MARTINET, *Economie des changements phonétiques*, Berne 1955, en particulier le chapitre 12, « Le dévoisement des sifflantes en espagnol ».

deux langues, il faut nécessairement faire intervenir le nombre des locuteurs de l'une et de l'autre : le latin, sous sa forme romane, était la langue de culture aussi bien à Saint-Gall ou à Reichenau, où l'on parle allemand aujourd'hui, qu'à Paris ou à Soissons. Mais il n'était pas autre chose à l'est, alors qu'à l'ouest, il était également la langue quotidienne de la majorité.

Un des facteurs les plus décisifs en la matière est la répartition des sexes chez les arrivants : si ceux-ci arrivent avec leurs femmes et, ce qui va normalement de pair, avec leurs familles et leurs possessions, il y a toute chance pour que leurs descendants continuent, au moins pendant un temps, à parler leur langue. Si, au contraire, les hommes arrivent seuls, bien décidés à trouver sur place les moyens de subvenir à leurs besoins de tous ordres, sexuels y compris, il est assez vraisemblable que les enfants issus des nouveaux contacts parleront la langue de leurs mères plutôt que celle de leurs pères, plus occupés à asseoir leur domination sur leur nouvel habitat qu'à intervenir dans l'éducation de leur progéniture. C'est en 911 que les pillards normands, de langue danoise, s'installent officiellement en Neustrie. Mais, en 1066, lorsque Guillaume traverse la Manche pour faire valoir certains droits sur la couronne d'Angleterre, ses compagnons sont francophones.

En face du cas de la Normandie, considérons celui du Périgord. Ici encore, le nom de la province renseigne sur l'origine du peuplement : la Normandie, c'est le pays des hommes (danois *mand* « homme ») du nord (danois *nord*) ; le Périgord, c'est la région où s'établissent quatre (gaulois *petru-*) peuplades en marche (gaulois *-corii*). Nous verrons plus loin que *petru-* est l'équivalent celtique régulier du latin *quattuor*, d'où le français *quatre ;* quant à *corii*, pluriel de *corios*, dont le radical se retrouve sous la forme *Heer* « armée » en allemand, il désigne, sans doute, non seulement les guerriers, mais l'ensemble du groupe humain qui se déplace avec familles, armes et bagages, dans l'intention d'occuper un nouveau territoire. Cette fois-ci, il n'est pas vraisemblable que ces gens abandonnent leur langue si l'opération est réussie. Les autochtones seront peut-être contournés s'ils se révèlent peu enclins à se soumettre, mais plus vraisemblablement éliminés ou réduits en esclavage. Les femmes seront sans doute épargnées, par galanterie, si l'on veut bien retenir les deux sens de ce terme. Mais leurs enfants, de statut servile, apprendront la langue des vainqueurs.

Le déplacement de la tribu en masse peut être déterminé par la pression exercée par une autre population. Les Celtes qui s'établissent dans ce que les Romains appelleront la Gaule semblent le faire souvent sous la pression des Germains. Les Volques, par exemple, qu'on trouve plus tard dans le Languedoc, entre le Rhône et Toulouse,

étaient, à date plus ancienne, établis dans l'Allemagne moyenne au contact immédiat des Germains. Ceux-ci ont utilisé leur nom, *wolkai*, comme désignation générale de leurs voisins méridionaux. Devenu régulièrement *walh-*, ce terme a été emprunté plus tard par les Slaves pour désigner les Celtes romanisés et les romanophones en général établis dans la péninsule balkanique. Les formes slaves sont *vlach* et *valach-*, la première désignant les bergers nomades de Yougoslavie, la seconde apparaissant dans le nom de la Valachie. En germanique, c'est un adjectif dérivé qui a survécu : *walh-isk-*, d'où *Welsh*, *welsch*. Le mot désigne en anglais les Gallois celtophones, en allemand les Suisses de langue romane.

L'arrivée des Huns en Europe, aux v[e] et vi[e] siècles, a été un facteur déterminant de la poussée des Germains vers l'Occident. Il nous faut renoncer à poursuivre la recherche qui nous permettrait de savoir dans quelles circonstances les Huns ont pris la route de l'ouest. On envisage une défaite qui leur aurait été infligée par les Chinois. Pour y voir clair, il faudrait déterminer quelle technique nouvelle a, dans chaque conflit, assuré la supériorité d'une armée sur l'autre, l'utilisation du métal, de nouveaux alliages, la monte du cheval, l'invention du char, celle de l'étrier qui permet au cavalier de frapper d'estoc, l'arc, la fronde, l'arbalète, la poudre, le char d'assaut et la bombe atomique, sans parler des techniques de défense qui répondent à chaque innovation dans l'agression.

Une des causes premières des mouvements de population est la catastrophe naturelle, sécheresse, ras-de-marée, cyclone qui tarit les sources de subsistance. Sur un rythme plus lent, il faut envisager les procès de refroidissement ou de réchauffement de l'atmosphère qui déterminent un renouvellement de la faune et de la flore. On pense, par exemple, que le réchauffement post-glaciaire a repoussé vers le nord les troupeaux de rennes qui ont dû longtemps représenter un des piliers de l'économie de l'Europe moyenne. Les hommes ont alors suivi le renne, remplacés par des populations plus méridionales désormais à leur aise dans les latitudes plus élevées.

Un des cas les mieux documentés de mouvement de peuple est celui des Cimbres et des Teutons. Ils étaient établis dans le Jutland, cette péninsule qui est une avancée de la plaine de l'Europe du nord dans la direction de la Norvège et qui sépare la mer du Nord, à l'ouest, du Cattegat, à l'est, voie d'accès maritime vers les eaux calmes de la Baltique. La mer du Nord est dure, souvent violente, et menace sans cesse les bas pays qui la bordent vers le sud-est. Vers la fin du ii[e] siècle avant notre ère, elle envahit le nord du Jutland, laissant, lorsqu'elle se retire, un chenal, l'Isefjord d'aujourd'hui, qui en isole la partie septentrionale. Les Cimbres se trouvaient dans ce qui se nomme aujourd'hui le Himmerland, au nord d'Aarhus (**kimbr-* est passé



Errance des Cimbres et des Teutons de –115 à –101

régulièrement à *himmer-*). Les Teutons occupaient sans doute un espace plus vaste, mais ceux qu'affecte le ras-de-marée étaient situés plus à l'ouest ; sur l'Isefjord, la ville de Thisted, où *Thi-* serait ce qui reste de *Teuton*, pourrait témoigner encore de leur ancienne présence en ces lieux. Leur territoire dévasté ne laisse à ces gens que le choix entre la mort par inanition et le départ à la recherche de nouvelles terres. Ils se dirigent vers le sud-est. Arrivés aux confins de la Bohême, ils se heurtent aux Boïens, une tribu celtique qui les rejette. Leur route s'incurve vers le sud. Ils pénètrent dans ce que les Anciens appelaient le Norique, c'est-à-dire la Styrie et la Carinthie d'aujourd'hui, au sud-ouest de Vienne. Ils y trouvent sur leur chemin une légion romaine qu'ils défont, en − 113. Cependant ils ne poursuivent pas vers le sud, mais tournent vers l'ouest. Aux approches de la Forêt Noire, ils rencontrent les Helvètes, des Celtes qu'on retrouvera un peu plus tard dans ce qui est la Suisse d'aujourd'hui. Certains d'entre ces Helvètes, les Ambrons, se joignent aux migrants qui pénètrent en Gaule, arrivent jusque dans ce qui sera la Provence, y rencontrent les Romains qu'ils battent à Orange, en − 109. Ils se dirigent alors vers l'Espagne. Ils s'y heurtent aux populations locales. Ils n'insistent pas et repassent en Gaule où ils se séparent. Les Teutons restent dans la Province, alors que les Cimbres retournent vers l'est pour gagner l'Italie par la voie classique des invasions, le col du Brenner. Les Romains, deux fois échaudés, se ressaisissent et envoient, contre les envahisseurs, leur grand général, Marius, qui, en − 102, rencontre les Teutons à l'est d'Aix-en-Provence, au pied de la montagne de Sainte-Victoire, et les écrase. Au dire des vainqueurs, c'est un carnage, les Teutonnes immolant leurs enfants avant de se donner la mort. L'année suivante, Marius va à la recherche des Cimbres qu'il anéantit auprès de Verceil (Vercelli) dans ce qui est aujourd'hui le Piémont. Ainsi se termine ce dramatique épisode qui illustre bien les conséquences d'un cataclysme naturel de grande ampleur. Est-ce l'énergie du désespoir qui a permis à ces Germains, qui ne disposaient pas sans doute de toutes les ressources de leurs adversaires, de défaire par deux fois les légions avant de succomber finalement devant un grand homme de guerre ?

La conquête du monde occidental par les « Indo-Européens » a dû être ponctuée par de tels conflits entre tribus et peuples de même origine lointaine. On voit d'ailleurs que l'apparentement linguistique des Russes et des Afghans ne fait pas, de leur conflit d'aujourd'hui, autre chose que celui qui a, naguère, opposé les Américains aux Vietnamiens.

En face des déplacements en masse des tribus et des peuples, il y a ce que nous avons déjà rencontré dans le cas des Normands : le départ, à la belle saison, de jeunes guerriers bien décidés à se tailler un

domaine dans les espaces occupés par des populations aux techniques guerrières moins raffinées. Dans le cadre du régime patriarcal qu'on suppose pour les anciennes sociétés de langue indo-européenne[3], les cadets, au sens le plus large du terme, ne pouvaient espérer échapper à l'autorité de l'aîné, chef du clan, qu'en prenant le large et en faisant confiance, pour leur subsistance, à leurs talents de prédateurs. Ces expéditions de jeunes conquérants avaient, dans l'Italie ancienne, reçu la consécration de la religion. Ce qu'on désignait en latin comme le *uer sacrum*, le printemps (mais comment traduire *sacrum ?*) était l'autorisation, voire la recommandation, de sacrifier aux divinités tout ce qu'on rencontrait sur son passage. On sait que, dans de telles occurrences, les divinités ont bon dos : les bêtes immolées servent en fait à la nourriture de ceux qui les immolent et les hommes mis à mort laissent des femmes dont les instruments du destin pourront faire leur profit. Ainsi apparaissent de nouveaux clans où naîtront de nouveaux cadets prêts à renouveler l'opération. De proche en proche, un pays peut être ainsi subjugué et, contrairement à ce qui s'est produit en Normandie, la langue des conquérants peut finalement l'emporter, surtout si ceux-ci retournent chercher des épouses « légitimes » dans les tribus plus anciennement établies. Qu'on pense au rapt des Sabines. Ces pratiques semblent avoir joué un rôle important dans le peuplement indo-européen des parties de l'Italie non encore occupées par des peuples puissants comme les Etrusques. Une première vague aboutit aux établissements latins au sud de l'Etrurie. Une deuxième, qui contourne, par les montagnes, les conquérants déjà établis, occupe l'Ombrie, puis la Sabine et déborde, avec les Samnites, sur la Campanie et le sud de la botte.

Tout ceci suggère que l'indo-européanisation, de l'Inde à l'Irlande et à l'Islande, a dû se faire de façon assez graduelle, de premières vagues de conquérants se noyant finalement dans les populations autochtones, ou, au contraire, formant des noyaux durs susceptibles d'être rejoints par de nouveaux arrivants. On pense aux Celtes d'Espagne, dits Celtibères. Ils sont séparés de leurs cousins de Gaule et de Grande-Bretagne par la chaîne des Pyrénées et les Euskariens d'Aquitaine. Ils sont assez puissants pour opposer aux Romains une résistance opiniâtre. Il faudra, en − 133, le siège et le massacre de Numance pour les réduire.

Là où c'est une langue indo-européenne qui l'emporte, on peut être sûr que la population sera métissée, les conquérants eux-mêmes provenant, selon toute vraisemblance, de divers mélanges plus anciens. Il n'est pas certain que l'unification se fasse, entre les diverses

3. Cf. ci-dessous, chap. XI, p. 232.

composantes de la population, dès les premières générations. Pendant des siècles, on pourra, sur place, distinguer entre les seigneurs ou les patriciens d'une part, les esclaves, les serfs ou la plèbe de l'autre. Mais, à plus ou moins longue échéance, les oppositions s'atténueront pour finalement disparaître ou ne plus exister que sous la forme d'une minorité méprisée dont on ne sait plus ce qui la distingue du reste de la population sinon la discrimination qu'on lui témoigne. C'est le cas des cagots du Béarn et du Pays Basque et celui des etas au Japon. A une plus grande échelle, on pense aux castes de l'Inde avec des intouchables au bas de la hiérarchie sociale.

Linguistiquement, l'absorption définitive d'anciennes populations subjuguées peut se manifester par l'apparition de nouvelles normes : on ne constate pas, dans les usages quotidiens, le remplacement d'une forme linguistique par une autre ; il y a beau temps que les habitudes linguistiques des classes inférieures ont influencé l'usage général. Mais ce qui était auparavant la marque d'un parler familier ou négligé finit par s'imposer comme la forme de la langue qu'on emploiera désormais dans toutes les circonstances de la vie, y compris dans la littérature et les conjonctures solennelles. En Espagne, au cours du XVI[e] siècle, au moment où, après l'expulsion des Maures, la conquête des Amériques scelle l'unité de la nation, la norme phonologique va basculer. L'ancienne, encore représentée, vers 1500, dans la grammaire de Nebrija, va céder la place à une autre qui intègre des confusions et des évolutions phoniques qui trouvent leur source dans les imitations imparfaites du roman d'Espagne par des sujets de langue basque, bien des siècles plus tôt[4] : au début du siècle, le bon ton voulait qu'on prononçât *viejo* « vieux » en donnant aux consonnes leur valeur française. Cent ans plus tard, on propose aux étrangers les prononciations d'aujourd'hui, avec un [b] initial et la *jota* « gutturale » que les Parisiens entendent comme leur *r* dévoisé.

La graphie de l'irlandais contemporain reproduit assez fidèlement celle des textes moyenâgeux qui remonte jusqu'aux gloses du VIII[e] siècle. Ces gloses présentent la langue utilisée alors par les moines irlandais, c'est-à-dire une forme très « usée » du celtique occidental qui a perdu une syllabe sur deux. Deux siècles auparavant, les inscriptions ogamiques de Grande-Bretagne nous offraient la même langue, mais avec toutes les syllabes que suggère la comparaison avec les autres branches de l'indo-européen et qu'on retrouve, beaucoup plus tôt, dans les quelques inscriptions gauloises qui nous restent. Ce n'est pas en deux cents ans que la langue a pu changer à ce point. En fait, les inscriptions ogamiques sont rédigées selon une norme qui, dès

4. Voir André MARTINET, *Economie...*, *ibid.*

l'époque, ne devait plus correspondre aux usages quotidiens. Ceux-ci devaient déjà refléter les déformations que les autochtones, les Fomores, les gens que la légende fait venir de « sous-la-mer », avaient fait subir à la langue des conquérants indo-européens.

CHAPITRE III

CONDITIONS GÉNÉRALES DE L'EXPANSION INDO-EUROPÉENNE

Ce qu'il faut garder en tête si l'on veut saisir la nature des rapports entre les différentes branches de l'indo-européen dans la pré-histoire et au début des périodes historiques, c'est que le peuplement de l'Europe d'alors n'avait aucun rapport avec ce que nous connaissons à l'époque contemporaine. Aujourd'hui, chaque individu se rattache à un état-nation. Il y a des frontières précises. A l'intérieur de ces frontières, la distribution des langues et des ethnies est de type très variable, mais l'existence de ces frontières nous donne l'impression d'une certaine simplicité d'organisation.

Cueillette et agriculture.

Pour l'Europe du 2ᵉ millénaire avant notre ère, la situation a dû être tout autre. Il y a partout des solutions de continuité, c'est-à-dire qu'une population déterminée s'étend sur un domaine assez vaste avec des espaces non occupés ou occupés par d'autres populations. Il s'agit encore fréquemment de nomades et, par conséquent, les contacts peuvent se relâcher sur un point et être rétablis ultérieurement. La densité de population est très faible. C'est encore une densité qui rappelle celle des époques pré-néolithiques qui ont précédé l'invention de la culture du sol. « Néolithique » est un terme très ambigu. Etymologiquement, le mot évoque une « pierre nouvelle », c'est-à-dire la pierre polie. Mais, dans l'usage actuel, beaucoup plus que la nature des outils de pierre, « néolithique » implique la culture du sol, et l'apparition du métal ne mettra pas fin à ce stade de l'évolution de l'humanité. Avant le néolithique, il faut supposer des sociétés qui vivent de la chasse ou de la cueillette, c'est-à-dire de l'utilisation des produits du sol, mais uniquement des produits naturels, ce qu'on appelait autrefois les « racines ». Nous trouvons encore, en Afrique du

Sud, les Boschimans qui vivent de cueillette. L'essentiel de leur alimentation est assuré par les plantes qu'ils ramassent ou déterrent et dont ils mangent soit les feuilles, soit les tiges, soit la racine. Parallèlement, ils pratiquent la chasse et en utilisent le produit pour compléter leur alimentation. Lorsqu'on dit « néolithique », on implique que ceux à qui on applique cette épithète n'en sont plus là. Non qu'ils aient tout à fait abandonné la cueillette et la chasse — la chasse, on ne l'abandonnera jamais —, mais une partie importante, puis essentielle, de la nourriture va être fournie par les produits de la culture du sol.

La culture du sol entraîne un accroissement considérable de la densité de la population. Pour nourrir une population au moyen de la cueillette et de la chasse, il faut une très vaste étendue : les végétaux utilisables pour l'alimentation de l'homme apparaissent naturellement assez espacés. Qu'on pense, par exemple, à la cueillette des champignons dans un bois. Au contraire, lorsqu'on a semé du blé sur un espace dégagé au préalable, la densité des tiges porteuses d'épi est très élevée. Le gibier lui-même dépend, pour sa subsistance, des produits naturels, végétaux ou animaux, ce qui limite son développement.

Au troisième ou au quatrième millénaire avant notre ère, à l'époque où il s'étend, le mode de vie néolithique est bien établi dans l'Europe danubienne. Ailleurs, il gagne du terrain, mais lentement, et la densité de la population doit rester très faible. Tant qu'on pratique très généralement la cueillette et la chasse, on peut estimer la densité de la population à 0,5 habitant au kilomètre carré. Or actuellement on trouve en France près de 100 personnes dans le même espace, avec des densités supérieures dans d'autres régions de l'Europe occidentale. Il faut donc envisager, dans cet espace vital, à l'époque qui nous intéresse, des solutions de continuité et un semi-nomadisme assez général.

Dans la mesure où « néolithique » veut dire culture de la terre, nous n'en sommes pas sortis. Bien plus, la cueillette existe encore — qui n'a cueilli des mûres sur les buissons ? — et la chasse également, même s'il faut importer le gibier au préalable. Il est intéressant de relever que la chasse est restée une occupation noble, interdite aux « gens de peu », en face de l'agriculture, longtemps considérée comme le fait de manants.

L'agriculture semblerait impliquer sédentarisation. Mais cela ne se produit pas automatiquement : la culture, très souvent, consiste en un jardinage qui ne s'étend pas très loin de la maison. D'autre part, la « maison » n'est pas un édifice permanent. Ce peut être une tente, une « yourte », comme chez les Mongols. Une fois la récolte faite, on se déplace. La raison en est simple. On constate que, dans une terre vierge, les plantes poussent beaucoup mieux que dans une terre

précédemment cultivée. Aujourd'hui, on a la possibilité de cultiver n'importe quelle terre parce qu'on a des engrais. Ce qu'on demande à une terre cultivable, c'est d'être plate. Les Landes étaient un pays presque désert autrefois parce que la terre ne se prêtait pas à la culture. On a commencé par y faire croître des pins. Mais aujourd'hui, on y rencontre de superbes moissons de maïs là où toute la terre a été renouvelée. L'existence d'engrais chimiques a donc changé complètement les données du problème. Mais au départ, lorsque s'installe le néolithique, les gens s'établissent pour une saison en un point où ils font pousser leur récolte. L'année suivante, ils se déplacent, parce qu'ils savent qu'ailleurs leur culture, en terre vierge, leur rapportera beaucoup plus. On peut le constater aujourd'hui encore dans la région parisienne : si vous labourez un morceau de pré où, depuis cent ans, il n'y a pas eu de culture, vous pouvez compter la première année sur une récolte miracle.

La culture de la terre n'élimine donc pas le nomadisme, puisque les gens vont se déplacer pour trouver de nouvelles terres. Plus tard, lorsque, la population augmentant, les gens se sont serré les coudes, on a découvert qu'on pouvait rester sur la même terre à condition d'utiliser la fumure, faite au départ avec le fumier des animaux domestiques et les excréments humains, mais aussi, beaucoup plus tard, avec des produits importés et, finalement, des produits chimiques. On a découvert l'assolement triennal lorsqu'on a pris conscience que les différentes plantes ne mettaient pas à profit les mêmes niveaux du sol, certaines plantes n'utilisant que les couches superficielles, d'autres des couches profondes. On pouvait donc, une première année, semer du blé, aux racines très superficielles ; la deuxième année, des betteraves qui mettaient à profit des couches plus profondes du sol ; la troisième année, on laissait la terre se reposer, ce qu'on appelait la jachère. Mais tous ces raffinements n'existaient pas aux époques qui nous intéressent.

Ce qu'il faut également garder en tête, c'est que, lorsque se produit l'expansion des peuples de langues indo-européennes, l'avance ne se fait pas sur un front. Ces peuples ont quelques chances de se trouver isolés les uns des autres. Lorsque la poussée se fait dans des zones de faible densité, les gens de même origine vont se trouver séparés par des zones peuplées d'autochtones, par des déserts, par des marais. Tant qu'il y a nomadisme ou semi-nomadisme, les gens de même origine pourront aisément perdre contact.

L'expansion celtique.

Comme illustration à grande échelle, prenons le cas de l'expansion des Celtes. Les Celtes, au moment où on les identifie, se placent en

Allemagne, en Thuringe et au Sud de la Thuringe. Ils occupent également la Bohême et ce qu'on désigne aujourd'hui comme l'Autriche. Ultérieurement, petit à petit, ils pénètrent dans ce qu'on a plus tard appelé la Gaule mais, au moment où les Romains arrivent, il s'en faut que toute la Gaule soit celtisée, voire sous la domination des Celtes. Quand les Romains y occupent la côte méditerranéenne et y créent la *Provincia romana*, la Provence, cette région n'est occupée par les Celtes que depuis deux siècles au plus et combien superficiellement, et il n'est pas dit que les Ibères aient été totalement délogés des plaines du Languedoc. En Gascogne, seules des franges, le Médoc, Bordeaux, Arcachon, ont été atteintes, avec peut-être une tête de pont à Toulouse. Ailleurs, l'espace est habité par les Aquitains de langue euskarienne, pratiquement le basque. Mais, à travers cette Aquitaine et les Pyrénées, plus au sud, les Celtes ont dû établir des points d'appui qui leur permettent de maintenir les contacts avec l'Espagne où certains d'entre eux sont fixés. Ce sont ceux qu'on désigne comme les Celtibères, non pas nécessairement parce qu'ils étaient un mélange de Celtes et d'Ibères, mais simplement, peut-être, parce que les Romains pénétrant en Espagne avaient eu à traverser des régions occupées par les Ibères pour entrer en contact avec eux.

Dans ce qui est aujourd'hui la Provence, on pose, à l'arrivée des Romains, deux siècles avant notre ère, l'existence de « Celto-Ligures ». Mais il s'agit de Ligures, selon toute vraisemblance de langue non indo-européenne, qui étaient sur place lorsque les Grecs, vers – 600, ont fondé Marseille. C'est deux siècles plus tard que, du nord, sont arrivés les Celtes qui ne semblent pas avoir occupé la côte, ponctuée par des établissements grecs. A l'intérieur, si l'on en croit l'archéologie, les Celtes ont cruellement imposé leur domination aux populations locales. Dans le Languedoc, si Narbonne évoque les Celtes (*-bona* « port », comme dans *Vindobona*, Vienne, « le port blanc » sur le Danube), le site d'Ensérune, près de Béziers, témoigne de l'importance des établissements ibères antérieurement à l'arrivée des Volques celtiques.

Ce n'est pas le long de la côte méditerranéenne que les Celtes ont pénétré en Espagne, mais à travers les Pyrénées, probablement par le col très accessible du Somport. En témoigne le site de Berdún situé au débouché sud de la route de ce col. Berdún, bourg perché sur une acropole, est, naturellement, avec la confusion castillanne de *b* et de *v*, un Verdun. L'étymologie de ce nom est évidente : *Ver-* provient de *wer-*, lui-même issu de **uper* du fait de la chute du *p* qui, nous le verrons plus loin, est le trait le plus caractéristique des langues celtiques. Il s'agit d'un adverbe indo-européen bien connu, qui peut comporter un *s-* initial, comme dans le latin *super*, mais qui peut en être dépourvu, comme l'allemand *über* et l'anglais *over*. Le grec *huper*

est ambigu parce qu'en attique tous les *u-* initiaux ont reçu une aspiration, qu'ils proviennent de *u-* ou de *su-*. Le sens de *wer-* est celui de « sur », « au-dessus », mais également celui de *super-* ou *hyper-* dans nos composés modernes du type *supermarché, hypermarché*. C'est lui qu'on trouve à l'initiale de *Vercingétorix* « le chef suprême des guerriers », un titre plutôt qu'un nom véritable. Quant à *-dūn*, avec un *u* long, c'est semble-t-il, au départ, une enceinte ; l'équivalent exact allemand est *Zaun* « haie » ; en néerlandais *tuin* est le jardin ; en anglais *town* est la ville. Ces variantes sémantiques rappellent celles qu'on trouve pour un autre terme, latin *hortus* « jardin », grec *khórtos* « enclos », angl. *yard* « cour », les dérivés angl. *garden*, all. *Garten*, français *jardin*, vieux slave *gradŭ* (russe *gorod*) « ville », emprunt probable au germanique (i.-e. **ghortó*). Un *verdun*, c'est donc initialement un site fortifié au sommet d'une colline, ultérieurement, peut-être, un site fortifié quelconque. Il est vraisemblable que, comme le suggère la toponymie, de telles forteresses ponctuaient la route qui reliait les établissements celtiques de Gaule et d'Espagne.

Plus au nord, que ce soit sur le Continent ou dans les îles, outre-Manche, la domination celtique était certainement plus ancienne, mais elle résulte, comme nous le verrons ci-dessous, de vagues successives et les Gaëls qu'on retrouvera plus tard en Irlande sont attestés, dans les premiers siècles de notre ère, en Grande-Bretagne, à l'ouest des Bretons. Il est indispensable de préserver une vision dynamique de ces peuplements, avec maintien, au début, des langues des autochtones. Au fur et à mesure que ces derniers apprendront la langue des envahisseurs, celle-ci va s'adapter à leurs habitudes et à leurs besoins. Les langues celtiques d'aujourd'hui présentent des traits de tous ordres, syntaxiques notamment, comme la fréquence des constructions du type *c'est... qui*, dont il n'y a pas trace dans les attestations les plus anciennes. Ces traits sont souvent partagés par les parlers romans d'Occident : on pense à la fréquence de *c'est... qui* en français, ou à la distinction de deux verbes « être » en castillan, qui se retrouve en irlandais et n'est pas tout à fait étrangère au français (*était*, dérivé de *stabat*, suppléant, dans la conjugaison du verbe *être*, les racines *es-* et *fu-* du latin).

La hiérarchie sociale.

Il est clair que presque toujours l'expansion indo-européenne se fait sous la forme de l'asservissement de populations locales par une aristocratie dominante. Pour utiliser un terme qu'ont repris les Allemands à l'appui de l'entreprise national-socialiste, les usagers de l'indo-européen se présentent au départ comme un *Herrenvolk*, un peuple de seigneurs. La nature des rapports entre la classe dominante



Les Celtes en Europe centrale et occidentale
au IIIe siècle avant notre ère.
Les noms des peuples celtes sont en *CAPITALES ;*
ceux des autres peuples en **minuscules grasses.**

et le reste de la population a pu être assez variable. On peut dans certains cas supposer l'existence de strates hiérarchisées, comme on les trouve à Sparte où, soumis à une population assez restreinte de guerriers, les Spartiates proprement dits, vivent, dans les faubourgs, des gens dépourvus de tous droits civiques, les périèques (*perioïkoi* « ceux qui habitent autour »), producteurs de biens permettant aux seigneurs de s'exercer aux arts martiaux. Une troisième strate, les hilotes, sont, purement et simplement, des esclaves d'état, corvéables et sacrifiables à merci. Selon toute vraisemblance, les périèques sont des Achéens, descendant de la première vague indo-européenne, arrivée en Grèce vers – 2000, ceux auxquels on doit la civilisation mycénienne. Les hilotes représentent peut-être des populations pré-achéennes. Les Spartiates sont des Doriens, les Indo-Européens de la seconde vague, arrivés en Grèce vers – 1200, dont on ne doit pas croire qu'ils provenaient nécessairement de la même branche que les Achéens. Le dorien, comme dialecte du grec, pourrait n'être que de l'achéen infléchi dans son développement par ce qu'on a désigné comme un superstrat, c'est-à-dire la langue de maîtres éliminée finalement par celle des gens du commun.

On peut naturellement, dans toutes les situations de ce genre, se poser la question des proportions respectives des envahisseurs d'une part, des autochtones ou des déjà métissés d'autre part. Il est toutefois bien difficile d'y répondre. Dans le cas de Sparte où l'on est assez bien renseigné, les maîtres ont pu représenter moins d'un dixième du total. Mais ceci a certainement varié à travers le temps. Dans les appréciations qu'on a faites, a pu intervenir la fréquence du trait de blondeur. On sait que les Gaulois sont présentés par les Romains comme des blonds. Mais si l'on pense à la population française contemporaine qui, si elle s'est renouvelée, l'a été par l'arrivée de Germains qui devaient être aussi majoritairement blonds, on peut se demander pourquoi il y a en France une nette majorité de bruns. On pourrait être tenté d'interpréter la blondeur comme une sorte d'affection du système pileux déterminée par la proximité de la Baltique. C'est là une hypothèse aventurée, mais qui peut recevoir aujourd'hui encore un début de vérification. C'est un fait que si l'on passe un été au bord de la Baltique, en Scandinavie, on a quelque chance de voir ses cheveux blondir sur les bords. Quant aux yeux bleus, c'est encore autre chose : on sait que les yeux clairs ont une extension plus large que les cheveux blonds. On a, par exemple, des populations britanniques très brunes avec des yeux bleus. Pour en revenir aux cheveux blonds et à une influence de la Baltique, on rappellera que la transmission des caractères acquis n'est pas démontrée. Mais il est clair que le maintien de mutations peut expliquer qu'un habitat ait une influence sur les types physiques des individus. Il se trouve simplement que, dans le

monde, ce sont des peuples dont on peut penser qu'ils ont, à une certaine époque, séjourné dans le nord de l'Europe, qui présentent des cheveux clairs. En France, il y a plus de blonds dans le nord et l'est qu'ailleurs et il en va généralement ainsi partout en Europe, Russie comprise. Le fait que, dans l'histoire, les conquérants indo-européens sont souvent décrits comme blonds ne veut pas dire qu'ils viennent à coup sûr de la Baltique. Si un habitat est en cause, il peut avoir eu une extension beaucoup plus large, encore que vraisemblablement septentrionale.

Quoi qu'il en soit de l'origine de la blondeur, il n'est pas possible de la faire intervenir lorsqu'il s'agit de populations disparues pour lesquelles on dispose, au mieux, des appréciations sporadiques et impressionnistes de certains auteurs. On rappellera simplement que les peuples de langues indo-européennes de l'antiquité ont généralement valorisé la blondeur. C'est là un trait physique fréquemment attribué aux héros[1].

1. Voir, par exemple, ce qu'en dit Jean HAUDRY, dans Deucalion et Pyrrha, *Etudes indo-européennes*, Lyon, Université Jean-Moulin, p. 1-12.

CHAPITRE IV

DONNÉES LINGUISTIQUES ET DONNÉES ARCHÉOLOGIQUES

La comparaison des langues indo-européennes a, nous l'avons vu, amené à reconstituer, pour une époque antérieure à leur différenciation, la forme de certains mots et la valeur qu'ils pouvaient avoir alors : pour « père » on reconstruit la forme **pətēr* et une valeur qui n'est pas celle de « géniteur » mais celle de chef de clan. Ces opérations, lorsqu'elles sont bien conduites, devraient permettre de se faire une idée de la nature de la société qui parlait la langue, de ses conditions de vie et de certains traits de son habitat : le grand prêtre, à Rome, est désigné comme *pontifex*. Ce mot s'interprète immédiatement, en latin, comme « le faiseur de pont », comme *panifex* est « le faiseur de pain », le boulanger. Toutefois, la comparaison révèle que le radical *pont-* désigne au départ le chemin. Cela suggère que les ancêtres des Latins vivaient dans une région où les chemins avaient la forme de passerelles, probablement à travers des marais, et c'est ce qui a fait penser qu'ils s'identifiaient avec les occupants des terramares de l'Italie du nord.

On a donc très vite pensé à utiliser les données de la comparaison pour replacer dans l'espace et dans le temps la communauté primitive[1].

1. Sur tous les rapprochements linguistiques mentionnés ci-après, on peut se reporter aux dictionnaires étymologiques suivants :

Oscar BLOCH et Walter von WARTBURG, *Dictionnaire étymologique de la langue française*, Paris, 1932.

Emile BOISACQ, *Dictionnaire étymologique de la langue grecque*, 4e édition, Heidelberg, 1950.

Carl Darling BUCK, *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*, Chicago, 1949.

Joan COROMINAS, *Breve diccionario etimológico de la lengua castellana*, Madrid, 1961.

Albert DAUZAT, Jean DUBOIS, Henri MITTERRAND, *Nouveau dictionnaire étymologique*, Paris, 1964.

Mer et lac.

Soit, par exemple, les désignations d'étendues d'eau, de l'étang jusqu'à l'océan. En français, on distingue nettement la mer, qui s'étend à perte de vue, et le lac, relayé par l'étang, pour les nappes de moindres dimensions. *Mer* remonte à lat. *mare* qui s'oppose à *lacus* à peu près comme en français. Les formes attestées dans d'autres langues indo-européennes permettent de poser une base **mor* qui apparaît, le plus souvent, dans des dérivés variables. On la rencontre un peu partout, mais il s'en faut qu'elle désigne uniformément la mer proprement dite, y compris la haute mer. En germanique, par exemple, le mot, là où il est attesté, désigne plutôt un lac ou un étang. C'est le cas en néerlandais où, avant d'être asséchée, la « Mer de Haarlem » était d'une étendue si restreinte qu'on aurait hésité à la désigner comme un lac. L'anglais *mere*, archaïque et provincial, est attesté dans le nom *Windermere*. Il s'agit, à l'origine, d'un lac de dimensions très modestes situé dans ce qu'on nomme le *Lake District* au nord-ouest de l'Angleterre. L'emploi, en allemand, de *Meer* pour « mer » n'est pas le fait d'une population réellement en contact avec la réalité océanique. Avec le même sens, l'allemand connaît le mot *See* (*die See* « la mer », mais *der See* « le lac », chacun avec le genre de l'équivalent français, ce qui n'est peut-être pas un effet du hasard[2]). L'anglais a, pour la même notion, *sea* qui s'oppose à *lake*, comme *mer* à *lac*. Ce mot, strictement germanique, a éliminé **mor* en Scandinavie, mais il y apparaît, en quelque sorte, dévalué par rapport à l'anglais *sea* : *sø*, en danois, désigne le lac ou une mer très calme et s'oppose à un autre mot, *hav*, qui désigne la mer violente, celle qui monte et qui descend, à mettre probablement en rapport avec *hæve* « lever, hausser », angl. *heave*, all. *heben* : avec un premier élément d'orientation, respectivement « est » et « ouest », on a *Østersøen*, qui est la Baltique, et *Vesterhavet*, qui est notre Mer du Nord. Le gotique a, pour la mer, un dérivé à partir d'un *mari* ; il apparaît dans *marisaiws* en

---

A. Ernout et Antoine Meillet, *Dictionnaire étymologique de la langue latine*, 2e édition, Paris, 1939.
H. S. Falk et Alf Torp, *Norwegisch-dänisches etymologisches Wörterbuch*, Heidelberg, 1911.
Sigmund Feist, *Vergleichendes Wörterbuch der gotischen Sprache*, 3e éd., Leiden, 1939.
Elof Hellquist, *Svensk etymologisk ordbok*, Lund, 1922.
Ernst Klein, *A Comprehensive Etymological Dictionary of the English Language*, Amsterdam, 1966.
Friedrich Kluge, *Etymologisches Wörterbuch der deutschen Sprache*, Strasbourg, 1910.
Max Vasmer, *Russisches etymologisches Wörterbuch*, Heidelberg, 1955.

2. Sur d'autres effets du bilinguisme romano-germanique dans l'Empire franc, cf., ci-dessous, p. 94 et 153.

composition avec *saiws* (angl. *sea*, all. *See*) pour désigner le lac. Ce dernier s'emploie également seul dans ce sens.

En celtique, **mor* apparaît très anciennement, dans le gaulois *Aremorici* qui désigne les habitants de l'Armorique « ceux qui sont près de la mer », comme *Arelate* « Arles » était la ville « près de la plaine », c'est-à-dire la *Crau*, avec *are-* « près de » qui correspond au grec *pará* (par exemple, dans *parallèle* « à côté l'un de l'autre ») et *-late* équivalent du grec *platús*, l'un et l'autre avec chute régulière en celtique du *p-* initial. Dans les langues attestées plus récemment, *mor* s'est établi avec le sens de « mer » en breton et en gallois. Mais les Irlandais ont, à côté de *muir*, un autre mot pour l'océan ou la haute mer. En face on a, pour le lac, les formes *loch*, *loc'h* de l'irlandais, de l'écossais et du breton, à côté de breton *lenn*, gall. *llyn*.

En slave, la situation rappelle celle du latin, avec **mor-* (russe *more*) qui désigne la mer et un autre mot (russe *ozero*) qui désigne le lac. En baltique, on trouve des dérivés de **mor-*, mais on emploie normalement des formes nouvelles. Il en va a peu près de même pour les Indo-Iraniens qui n'ont de traces de **mor* que dans un dérivé. En grec, la seule forme pour « mer » est *thalassa*, *thalatta* qui, s'il est indo-européen d'origine, voudrait dire « la profonde », avec la base qu'on retrouve dans l'anglais *dale*, all. *Tal* « vallée » et le vieux-slave *dolŭ* « puits ». On entend, en anglais, *the deep* « la profonde », pour la mer. L'opposition est nette, en grec, entre *thalassa* et *limnē* « le lac » et « l'étang », ce dernier d'une base qui évoquerait plutôt la vase des marais que les vagues de l'océan.

Il faut enfin rappeler qu'à partir de la base **mor*, on a des formes parallèles ou dérivées qui désignent le marais. Le mot *marais* lui-même, comme l'anglais *marsh*, représente un germanique *mar-isk*. L'anglais *moor*, avec une ancienne voyelle longue, désignait au départ le marais.

Ce qu'on peut conclure de tout ceci, c'est que la forme la plus répandue, que nous avons désignée comme la base **mor*, devait désigner au départ une étendue d'eau sans spécification particulière, aussi bien une zone marécageuse qu'une nappe s'étendant à perte de vue. Cela semblerait indiquer qu'on ne connaissait pas la vie proprement maritime qui caractérise le contact avec la haute mer et qui entraîne une distinction nette entre lac et étang d'une part et mer d'autre part. Arrivés au contact d'une mer à marées de grande ampleur et de ressac violent, ce qui crée des conditions très particulières pour la navigation, il a été nécessaire d'établir des distinctions, soit en opposant **mor* ou ses dérivés à un mot pour lac, comme en latin et en slave, soit, comme en germanique, en opposant à **mor*, réduit au sens de lac ou d'étang, une nouvelle désignation de la mer, susceptible d'être à son tour détrônée par un concurrent pour renvoyer à la haute

mer, soit encore, comme en grec, en écartant **mor* tout à fait au profit de deux nouveaux venus.

Le hêtre.

Un vocable qui a été à la source de nombreuses spéculations est le nom du hêtre, en latin *fāgus,* mot qui subsiste en ancien français sous la forme de *fou,* d'où le diminutif *fouet* et le verbe *fouetter.* La forme correspondante est, en allemand, *Buche,* avec le même sens. On en a dérivé le mot qui désigne le livre, *das Buch,* et la lettre, *Buchstabe,* proprement « bâton de hêtre », dérivés l'un et l'autre de l'usage qu'on faisait du bois et de l'écorce de cet arbre pour graver des messages. Toutes les langues germaniques connaissent cette désignation du hêtre, parfois sous une forme un peu modifiée, comme l'angl. *beech* ou le danois *bøg.* Le mot germanique a été emprunté en slave, et le mot latin en celtique. En Grèce, où l'on ne trouve pas de hêtre, l'équivalent *phēgós* désignait une sorte de chêne. Certains chercheurs, en Allemagne, désireux de pousser vers l'ouest l'habitat primitif des Indo-Européens, ont fait valoir que le hêtre ne pousse pas au nord-est d'une ligne qui part de Königsberg, aujourd'hui Kaliningrad, pour se diriger vers la Crimée. A cela on a répondu en faisant remarquer que la désignation des noms d'arbres se fonde assez souvent sur les utilisations qu'on fait de l'arbre ou de ses fruits à différentes fins, plutôt que sur son identité botanique. La forme de base *fāgus, Buch, phēgós* pourrait être rapprochée du grec *phágein* « manger » et désigner en fait l'arbre qui produit ce qui a pu être, avant les céréales, une des bases de l'alimentation, la faine du hêtre ou le gland de certains chênes. Dans ce cas, l'origine et la diffusion du terme ne dépendraient plus de la présence du hêtre. Le rapprochement, moins évident, avec des formes slaves désignant le sureau serait justifié du fait des usages alimentaires des baies de cet arbre encore attestés en Scandinavie.

Le poisson.

Les conclusions que l'on pourrait vouloir tirer des désignations du poisson, dans les différentes langues, relativement à un habitat primitif illustrent bien le caractère hautement hypothétique de telles considérations. Au centre du domaine, du baltique au grec et à l'arménien, on a des formes dérivées d'une base première que les linguistes identifient bien, mais qui est susceptible de perdre son identité parce qu'assez courte et formée de phonèmes instables. Qu'on en juge par les produits suivants : grec *ikhthûs,* vieux-prussien /zuk/, en lituanien *žuvis,* mais également *žuk-* en composition, arménien *jukn.* On s'attendrait à ce que le slave soit représenté ici, ce qui

La charrue.

Si les différentes formes que nous venons d'examiner n'apportent guère de certitudes relatives à une ancienne localisation géographique d'un peuple indo-européen à l'époque de sa cohérence, on peut, par l'examen de certaines innovations, tirer des conclusions sur la chronologie relative de certains mouvements. La technique du labourage au moyen d'un instrument tracté se manifeste par la base que l'on rencontre dans *araire, aratoire.* On la retrouve dans toutes les langues européennes de la famille, ainsi qu'en arménien, mais elle est absente des langues de l'Inde et de l'Iran, ce qui peut faire supposer que ceux qui parlaient ces langues se sont séparés avant l'invention de la charrue. Ou doit-on envisager, pour ces peuples, une longue période de nomadisme qui aurait fait tomber dans l'oubli le vocabulaire afférent au labourage par traction animale ? Sans doute le vocabulaire peut-il se renouveler à l'occasion de renouvellement des techniques. C'est ce qui s'est passé lorsque la charrue, munie de roues, a remplacé l'ancienne araire, au départ simple branche coupée au-dessous d'un rameau latéral épointé pour pénétrer dans le sol. Les vieilles langues germaniques conservent la base *ara-* pour « labourer », mais utilisent déjà le néologisme angl. *plough, plow,* all. *Pflug* pour l'instrument. Un *p* initial en germanique indique presque à coup sûr un emprunt ou une création expressive.

Les données archéologiques.

Pour étayer ou compléter les données de la comparaison, on a, bien entendu, pensé à un recours aux données archéologiques[3]. Mais, pendant très longtemps, il est resté très difficile de combiner les apports des deux disciplines. Les archéologues se sont, à juste titre, beaucoup intéressés aux poteries, aux différentes techniques qu'elles manifestaient, à l'évolution de ces techniques à travers le temps et à la diffusion des différents types à travers l'espace. Malheureusement, il ressortait que cette diffusion se faisait le plus souvent du sud vers le nord, alors que les déplacements de peuples que suggéraient les données historiques ou les conclusions des comparatistes se faisaient plutôt du nord vers le sud. Cela voulait dire simplement qu'il existait

3. On s'est inspiré essentiellement ici des données présentées dans Edgar C. POLOME (dir.), *The Indo-Europeans in the fourth and third millennia*, Ann Arbor, 1982 ; d'autres sources de documentation sont représentées par Guy RACHET, *L'univers de l'archéologie, technique, histoire, bilan*, 2, Verviers, 1970 et André LEROI-GOURHAN et son équipe, *La préhistoire*, Paris, 1966. Les datations varient beaucoup, selon les auteurs, avec une tendance à reculer dans le temps.

des relations commerciales entre des peuples établis à différentes latitudes et qu'on ne pouvait, bien entendu, pas conclure de la présence, dans deux régions différentes, de poteries du même type à un apparentement linguistique des populations en cause.

Il en va certes autrement des types de sépultures. Si l'on constate qu'un type donné, parfaitement caractérisé, se retrouve dans des régions différentes avec des datations successives, il semblerait qu'on puisse souvent en conclure à un déplacement de population. Mais, là encore, une mode pourrait gagner de proche en proche. Reste à savoir quelle était la langue que pratiquaient ces gens. Les objets qui accompagnaient le défunt dans sa tombe avaient pu être le fruit d'échanges commerciaux de peuple à peuple. Certains traits physiques relevés dans les ossements, comme le rapport entre la longueur et la largeur du crâne, ont longtemps retenu l'attention. On a distingué entre des gens à crânes allongés, ou dolichocéphales, et des gens à crânes courts et larges, ou brachycéphales. En comparant les données de ce type dans les populations établies aujourd'hui et celles qu'on relevait sur les ossements dégagés, on pouvait esquisser des sens de déplacement. Mais pouvait-on attribuer la dolichocéphalie ou la brachycéphalie aux usagers de telle ou telle langue, sans recourir, au départ, à des hypothèses invérifiables ?

Avant de voir comment se pose aujourd'hui le problème des rapports entre l'archéologie et la localisation, dans le temps et l'espace, des populations indo-européennes, il peut être intéressant de résumer rapidement les données archéologiques relatives aux mouvements culturels en Europe occidentale et centrale, sans remonter, dans le passé, jusqu'aux âges glaciaires.

Le néolithique.

La première expansion dont on doit faire état est celle du néolithique. Il s'agit d'une culture qui n'est pas, comme son nom semble le suggérer, caractérisée par un nouveau traitement du silex, mais par le passage de la cueillette à l'agriculture. On notera toutefois que le métal n'est pas apparu et que les outils restent de pierre. Cette expansion résulte-t-elle d'une acculturation se propageant par imitation de proche en proche, ou de mouvements de populations, de tels mouvements étant le fait d'agriculteurs se déplaçant au fur et à mesure de l'épuisement des couches supérieures du sol ? On ne peut guère se prononcer à coup sûr et il faut propablement supposer l'un et l'autre. En Europe, le néolithique se manifeste dès le 7[e] millénaire, dans les Balkans et la plaine danubienne, en provenance, sans doute, du Proche-Orient. Il y paraît rayonner autour des sites danubiens où nous le retrouverons ci-dessous lorsque nous chercherons à évaluer sa

contribution fondamentale aux cultures européennes qui ont façonné le monde contemporain. De là, au 5[e] millénaire, il s'étend à l'Italie et, par la Provence, jusque dans la péninsule ibérique. Vers le nord, il gagne l'Allemagne et la Scandinavie puis, s'incurvant vers l'ouest, les Pays-Bas, la France du nord et la Grande-Bretagne.

C'est au cours du 4[e] millénaire que se complète l'occupation néolithique de l'Europe occidentale, que sa propagation se fasse par le nord ou par le sud. Mais cette culture, au moment même où elle s'établit là en profondeur, s'est déjà trouvée exposée depuis quelques siècles, dans son site danubien d'origine, à des pressions venues des steppes eurasiennes. Peut-être ces pressions sont-elles à la source de déplacements de population qui ont hâté et intensifié cette néolithisation de l'Ouest européen qui en a conditionné l'évolution économique ultérieure de façon beaucoup plus permanente que les vagues successives de conquérants indo-européens, qu'il s'agisse des Gaulois, des Romains ou des Francs.

Les mégalithes.

L'expansion, d'est en ouest, du néolithique va entrer en contact avec une poussée venue du sud et gagnant vers le nord le long des côtes, celle des mégalithes. Rien n'a laissé de traces aussi évidentes sur le terrain que ce qu'on peut désigner comme la culture mégalithique. Les mégalithes, ce sont les « grandes pierres », en breton *men hir*, qu'on trouve soit isolées, soit groupées en rangées comme à Carnac en Bretagne. On ne sait trop quelles en étaient les finalités. On pense, bien entendu, à un culte. Mais, à côté des *menhirs*, on trouve des « tables de pierre », les *dolmens*, initialement recouverts de terre et formant des caves qui ont servi comme lieux de sépulture, un peu comme les caveaux de famille contemporains. A noter que les corps qu'on y retrouve correspondent à des décès de dates différentes et non, selon toute vraisemblance, à des mises à mort destinées à l'accompagnement d'un puissant défunt. Le mouvement mégalithique, d'origine sans doute orientale, mais dont on voit mal comment il s'est propagé jusqu'en Occident, paraît débuter, en Europe, au 4[e] millénaire, en Espagne d'abord, puis en Bretagne et dans la France de l'ouest et dans les îles britanniques et, enfin, au troisième millénaire le long de la mer du nord, jusqu'au Danemark. S'agit-il d'un peuple qui se déplace ou d'un comportement qui se diffuse progressivement selon un processus que nous serions tentés d'assimiler à une évangélisation, auprès de gens parlant la même langue ou des langues différentes ? On est en peine de se prononcer. Les considérations développées ci-dessus au sujet de la désignation de la mer, jointes à bien d'autres, suggèrent qu'il ne s'agit pas, à cette date et dans ces régions, de populations de

langue indo-européenne. Il convient, semble-t-il, de bien distinguer entre cette poussée mégalithique et un renouveau ultérieur, celui auquel on doit l'ensemble de Stonehenge, en Angleterre, beaucoup plus savamment élaboré que les alignements de Carnac, par exemple.

Les kourganes.

Avant de retrouver, dans le bassin du Danube, les premiers représentants du néolithique européen et pour, finalement, mieux mettre en valeur l'originalité de leur culture, nous nous transportons jusque dans les steppes de l'Eurasie où l'archéologie contemporaine a identifié ce qu'on désigne comme la culture des kourganes. On a de bonnes raisons de penser qu'elle était celle d'un peuple qui parlait, au sixième millénaire, entre l'Oural et la Caspienne, une forme de l'indo-européen commun.

Les kourganes, en anglais *pit-graves*, sont des lieux de sépulture mi-enterrés, avec des parois et une couverture de pierre recouverts de terre. Ce sont les sépultures de chefs qu'on suppose riches et puissants si l'on en croit les trésors et le nombre des serviteurs et concubines qui les accompagnent dans la mort. Cela rappelle le sacrifice des veuves sur le bûcher du mari qu'on a signalé dans l'Inde presque jusqu'à nos jours. Les richesses que recèlent les kourganes ne semblent pas être le produit d'industries locales et ont quelques chances de provenir de rapines plutôt que de négoces. Tout indique l'existence d'une société très hiérarchisée, avec une classe dirigeante de chefs et également de prêtres, si l'on extrapole à partir de cultures de même type bien décrites, avec les brahmanes de l'Inde et les flamines de Rome. On trouve ensuite une classe de guerriers et enfin des pasteurs chargés d'assurer la survivance. Il faut toutefois se garder d'une identification trop poussée avec les structures sociales d'époques plus récentes touchées par la culture néolithique et les activités agricoles qui la caractérisent. La complexité de la structure sociale qu'impliquent les données archéologiques paraît exclure une économie fondée exclusivement sur la chasse et la cueillette, et c'est pourquoi on doit supposer que des pâtres guidaient leurs troupeaux à travers les steppes. A un certain point du temps, on va domestiquer le cheval, ce qui facilitera les activités prédatoires. Pour le faire participer à des actions proprement belliqueuses, il faudra sans doute attendre l'invention du char qui assure au guerrier une stabilité qu'il ne trouvera, à cheval, que beaucoup plus tard, avec l'étrier.

Les sépultures du même type que les kourganes, avec victimes immolées, vont ponctuer, à travers l'Europe, l'avancée des Indo-Européens. On en a trouvé en Bohême et en Allemagne moyenne, à Unětice par exemple avec une datation de – 1700. Elles sont attestées

dans le Caucase du nord, en Transcaucasie et en Anatolie sur le trajet probable de certaines vagues de conquérants indo-européens. Les datations n'indiquent nullement un point d'arrivée, car l'immolation des proches a dû se poursuivre à travers les âges, même si en l'absence de documentation archéologique nous ne pouvons en marquer les étapes. On la retrouve à l'époque mérovingienne dans la tombe de Childéric I[er] à Tournai. Les sacrifices humains sont bien connus dans l'antiquité classique, ne serait-ce que sous la forme spectaculaire des combats de gladiateurs. Les massacres à grande échelle des guerres contemporaines où l'on immole des millions d'êtres humains sur l'autel de la patrie, de la démocratie ou de telle autre abstraction nous rappellent la permanence de certains comportements qui ne nous révoltent que lorsque nous ne percevons plus les justifications qu'on en pouvait donner.

Les Danubiens

Les fouilles faites à date relativement récente dans les Balkans et dans le bassin du Danube permettent de mieux apercevoir les conditions dans lesquelles se sont produites les expansions indo-européennes vers le sud-ouest à partir des steppes de l'Eurasie. Dès le septième millénaire avant notre ère, se développe dans ce coin du monde, berceau du néolithique européen, une culture qui va être exposée, à partir du cinquième millénaire, aux incursions des nomades des kourganes et qui reculera de ce fait vers les côtes de la mer Egée et la Crète pour céder finalement vers la fin du troisième sous la pression d'Indo-Européens, les Achéens, mais non sans avoir profondément marqué les envahisseurs. Ce que l'archéologie révèle de cette culture la place aux antipodes de ce que devait être la structuration sociale des conquérants. Les néolithiques danubiens apparaissent comme les représentants d'une société matriarcale où la divinité centrale est une déesse mère de la fécondité et où les techniques agricoles et la production des objets de consommation paraissent plus en honneur que la pratique des arts martiaux. Les sites habités les plus anciens ne révèlent pas de réels dispositifs de défense contre l'agression extérieure. Ce n'est que tardivement, lorsque se fera sentir la pression récurrente de leurs turbulents voisins, qu'apparaîtront des embryons de fortifications qui finalement se révéleront inefficaces contre les raids de cavaliers puissamment armés.

Si l'on garde en tête l'ancienneté des manifestations de cette culture, on est étonné de l'avance qu'elle représente sur la voie du développement de l'humanité. On a longtemps pensé qu'au troisième millénaire avant notre ère, ce n'était qu'en Mésopotamie et dans la vallée du Nil qu'on devait chercher les premières traces, par exemple dans le

domaine de l'écriture, de ce qui devait déboucher sur la culture de l'Occident. En fait, on peut suivre chez les Danubiens l'évolution d'un graphisme à partir de signes d'origine cultuelle, qui mène, au quatrième millénaire, à ce qui paraît être un syllabaire, dont on peut se demander s'il ne serait pas à la source de ceux qu'on retrouve plus tard en Crète.

Un exposé où l'unité chronologique n'est plus, à grande échelle, le siècle, mais le millénaire ne met pas en valeur le rôle décisif qu'a dû jouer, dans le développement du sous-continent européen cette culture danubienne. Il convient d'insister sur le fait que, de son apparition à son déclin, elle a dû se prolonger pendant plus de deux mille ans, non sans laisser subsister l'essentiel de ses contributions aux progrès de l'humanité. On a souvent parlé du miracle grec. Même si nous prenons nos distances vis-à-vis d'un terme qui ne peut que bloquer la réflexion et freiner la recherche, nous pouvons y voir une façon un peu naïve de caractériser cet amalgame assez remarquable de puissance créatrice et de vigueur expansionnelle issu de la symbiose de la fécondité danubienne et de l'agressivité des nomades. Cette synthèse dont la Grèce classique nous paraît être le modèle est parfaitement illustrée par le syncrétisme de la religion hellénique où le panthéon masculin des conquérants indo-européens voisine avec les déesses de la fécondité et de l'intelligence. D'une part, le clan des mâles, avec à sa tête le dieu du ciel et du tonnerre, Zeus, et, comme pendants féminins, une Vénus, ou sa réplique nordique Freyja, conçue comme le « repos du guerrier », et une Héra, la Junon des Romains, épouse du chef et protectrice des héros, avec de part et d'autre, ce qui n'est pas fortuit, une base *hēr-*[4]. D'autre part, les divinités proprement helléniques de la terre, de la fécondité et de l'invention : Gê, Déméter, Perséphone et Athéna.

Cette symbiose n'existe proprement que chez ceux qui ont participé aux cultures du monde égéen. D'autres Indo-Européens, dont on pense qu'ils ont pu au cours de leurs déplacements séjourner quelque temps dans le bassin du Danube, ou bien, comme peut-être les Indo-Iraniens, ne participent nullement à cet amalgame, ou, comme les Latins, ne le connaissent qu'à titre d'emprunt tardif.

Première poussée des gens des steppes.

Vers la fin du cinquième millénaire, les gens des steppes pénètrent une première fois dans l'espace danubien. Leur base de départ paraît

4. Cette base paraît s'identifier avec celle d'angl. *year*, all. *Jahr*, « année », proprement « l'année nouvelle », le héros étant celui qui, au-delà de la mort hivernale, peut participer au renouveau annuel. Sur tout ceci, voir Jean Haudry, Héra 1, *Etudes indo-européennes* 6, 1983, p. 17-41, Héra 2, *ibid.* 7, 1983, p. 1-28, et Héra et les héros, *ibid.* 12, 1985, p. 1-51.

être les cours inférieurs du Don et du Dniepr. Plutôt que d'une invasion, il s'agit d'incursions de cavaliers qui imposent leur domination sur certains points du domaine. Les traces de ces agressions se retrouvent, depuis les bouches du Danube en remontant le fleuve jusque dans ce qui est aujourd'hui la plaine hongroise et, au sud, jusqu'en Macédoine. Les populations locales ont dû parfois opposer aux envahisseurs une résistance assez efficace. Cela a été le cas, semble-t-il, des premiers atteints, dans la région du Dniestr, dans ce qui est aujourd'hui l'Ukraine occidentale et la Moldavie. Avaient-ils une expérience antérieure des agressions de leurs voisins de l'est ? Dans la plaine de la Tisza, des populations ont été subjuguées et leur culture anéantie. D'autres se sont repliées vers le midi et l'Occident, non sans avoir dû établir un *modus vivendi* avec l'envahisseur. Il est naturellement impossible de préciser ce que pouvait être la situation linguistique dans les régions en cause à l'issue de ces mouvements. On peut simplement supposer que ces incursions n'ont pas amené un afflux de nouvelles populations et que, là même où les conquérants se sont finalement imposés, leur mélange avec les autochtones a pu aboutir à éliminer leur langue, un peu sur le modèle de ce qui s'est passé, à cinq mille ans de là, en Normandie.

Deuxième poussée des gens des steppes.

C'est près d'un millénaire plus tard que se produit une nouvelle poussée des gens des steppes qui va aboutir à l'occupation d'une grande partie de l'Europe par des populations de langue indo-européenne. Pas plus qu'au cours de la première vague, ne doit-on alors supposer une élimination des populations locales. Il y aura, le plus souvent, sans doute, amalgame et, linguistiquement, il n'est pas dit que l'indo-européen s'impose partout. Le point de départ de cette nouvelle vague semble devoir être situé plus loin à l'est que celui de la première, dans la région au nord du Caucase. Cette fois, la poussée se fait non seulement vers l'ouest et le sud-ouest, mais également vers le nord-ouest. Sont occupées ce qui est aujourd'hui l'Ukraine, la Pologne et l'Allemagne jusqu'à l'Elbe et au-delà. Est également atteinte l'Italie du nord et du centre[5]. Vers le sud, on ne dépasse guère la Macédoine, mais la poussée déborde sur l'Asie mineure. Dans cet espace, se produiront plus tard divers mouvements qui aboutiront finalement à une totale indo-européanisation. Ce n'est certainement pas cette deuxième vague qui a apporté en Italie ce qui sera le latin et les langues italiques.

5. D'où, peut-être, certaines gravures rupestres du Val Camonica ; cf. Isabelle TURCAN, *Etudes indo-européennes* 6, 1983, p. 1-15.

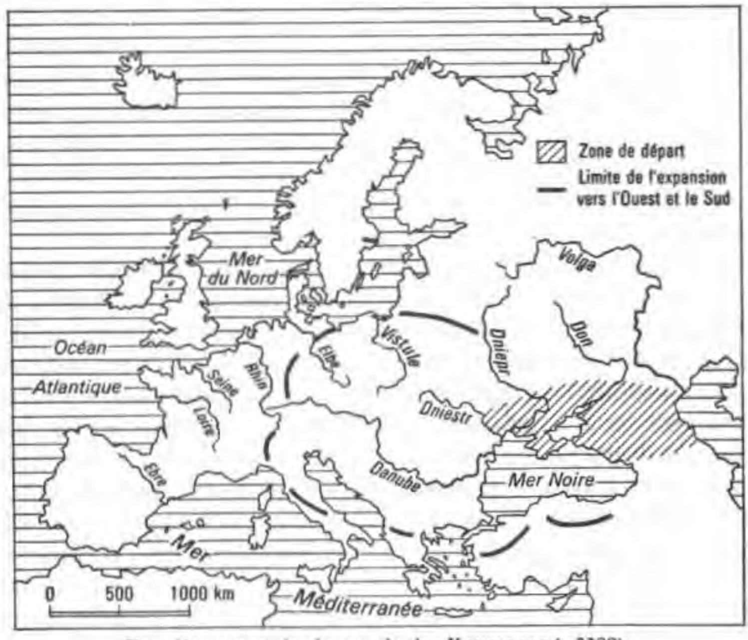
Troisième poussée des gens des steppes.

Plus tard, au début du troisième millénaire, une troisième vague d'envahisseurs des steppes, venue d'une zone s'étendant du Dniestr à l'Oural, recouvrira de nouveau l'Europe centrale. On supposera ici que l'indo-européen apporté par les dernières vagues dans l'espace danubien sera à la source, notamment du grec, d'une part, des langues aryennes, de l'autre, alors que, ce qu'on désignera comme l'indo-européen de l'ouest, italique, celtique et germanique, dérive de la langue des envahisseurs de l'Europe du nord. Dans cette zone septentrionale et à cette époque, on place une culture dite de l'amphore globulaire. Elle fera place un peu plus tard à la céramique cordée dont on trace la diffusion de la Roumanie, à travers la plaine du nord, jusque dans les Pays-Bas et la Scandinavie. Il s'agit d'une poterie caractérisée par des impressions réalisées dans la glaise au moyen de cordes. Elle s'accompagne de la hache de combat en cuivre ou en pierre polie sur le modèle des instruments de métal.

La diffusion vers l'ouest de la céramique cordée est contemporaine de l'expansion d'un autre type de poterie, connu sous le nom de gobelets campaniformes, qui paraît provenir d'Espagne. Il pourrait sans doute s'agir d'une expansion commerciale. Mais les archéologues tendraient à identifier sa poussée vers le nord-est avec celle d'un peuple d'archers se déplaçant parallèlement à la culture mégalithique, mais un peu plus à l'intérieur des terres. Cette activité des représentants des cultures de l'Occident explique sans doute le caractère tardif de l'expansion indo-européenne vers les côtes de la Mer du Nord et au-delà du Rhin. Entre la deuxième vague et la pénétration en masse des Celtes dans ce qui deviendra la Gaule, il s'écoulera plus de deux millénaires.

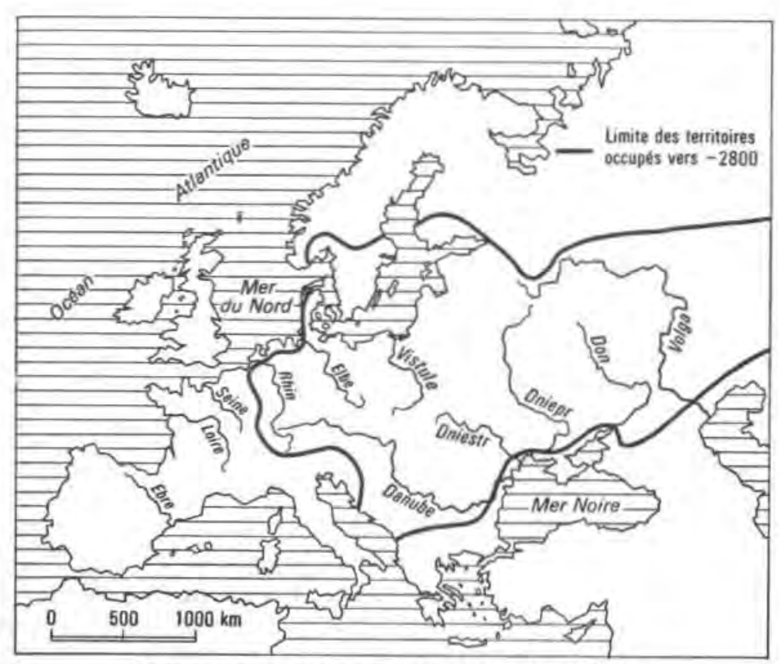
Vers la fin du troisième millénaire, on peut supposer des formes de l'indo-européen bien établies, de l'Allemagne moyenne à l'Ukraine, de la Baltique aux Balkans, dans un espace où elles restent certainement en concurrence avec des formes linguistiques plus anciennes, mais où elles représentent la langue des éléments les plus puissants et les plus dynamiques de la population. Sans doute, ne faut-il pas oublier que partout où l'agriculture reste soit inconnue, soit pratiquée de façon nomadique, c'est-à-dire sous forme de jardinage sans fixation permanente, la densité de la population reste très faible. L'arrivée, dans une région, d'éléments allogènes, n'exclut donc pas une coexistence parallèle, surtout si les moins forts se contentent des parties les moins fertiles et les moins giboyeuses du territoire. C'est à partir de ce domaine que vont probablement se diffuser les langues indo-européennes, à l'exception des langues anatoliennes. D'abord, peut-être,

Première poussée du peuple des Kourganes (−4200)



Deuxième poussée du peuple des Kourganes (−3300)

Cette poussée est contemporaine de l'extension des mégalithes le long de l'Atlantique et autour de la Mer du Nord



Troisième poussée du peuple des Kourganes (−2800)

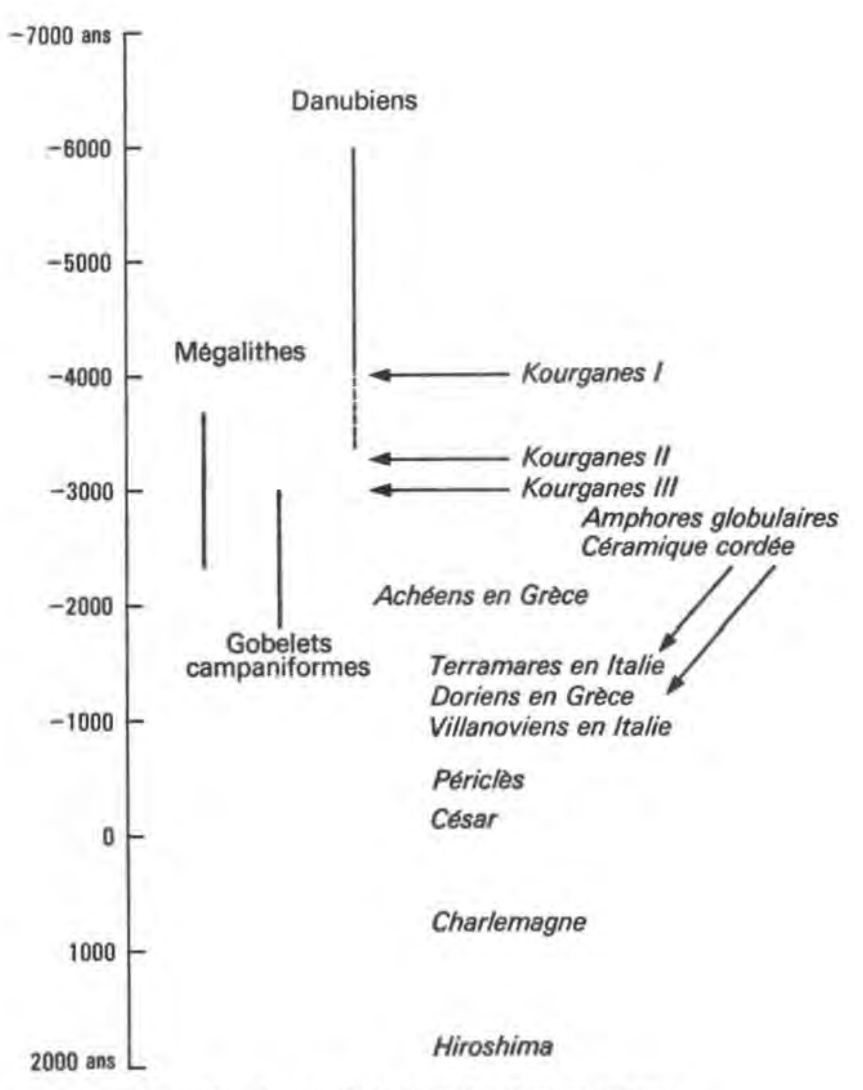
Par rapport à la deuxième poussée, on notera une moindre avancée vers le Sud et la progression dans les autres directions. Cette poussée est contemporaine de l'extension des archers porteurs des gobelets campaniformes de l'Espagne à la Bohême[6].

les langues aryennes, vers l'Asie mineure, puis le grec vers le sud ; le latin, puis les autres langues italiques, du nord vers le Danube, puis vers l'Italie ; les Celtes à leur suite, puis vers l'ouest, les Germains vers le nord d'abord, puis, sur les traces des Celtes, vers le sud.

A l'aube de l'histoire.

Au tournant du dernier millénaire avant notre ère, apparaît le rite funéraire de l'incinération avec enterrement des réceptacles contenant les cendres dans des cimetières qu'on désigne comme les champs d'urnes. On retrouve ces champs dans la plaine du nord, en Lusace, mais également en Italie septentrionale avec la culture villanovienne qui rayonne à partir de l'Emilie. Cette culture a été précédée, plusieurs siècles auparavant, dans des sites placés plus à l'ouest, par celle des

6. Ces cartes s'inspirent directement de celles de Marija GIMBUTAS, dans *The Indo-Europeans in the fourth and third millennia*, p. 53, 54 et 55.



Chronologie des cultures et des déplacements en Europe



L'Europe à l'aube de l'Histoire – Carte d'orientation

terramares. Plus au nord, on rencontre la culture dite atestine, celle des Vénètes. Toutes ces cultures sont attribuées à des peuples de langues indo-européennes. Les gens des terramares sont les premiers à descendre vers le sud et l'on peut voir en eux les ancêtres des Latins. Les Villanoviens, lorsque, contournant les Etrusques, ils pousseront le long de l'Appenin, seront identifiés comme les Ombriens, les Sabins, les Osques et les Samnites, appartenant tous à un groupe linguistiquement bien caractérisé. Les Vénètes, longtemps assimilés à des Illyriens, sont, par leur langue, à rapprocher des Latins. Culturellement on distingue mal les Villanoviens des Etrusques situés historiquement en Toscane, c'est-à-dire plus au sud. Mais ils s'opposent à eux par la langue, l'étrusque n'étant, de toute évidence, pas une langue indo-européenne.

C'est également vers la fin du deuxième millénaire qu'on place un mouvement, dit « des longues épées », parti du nord, à travers l'Europe jusqu'aux Balkans. On l'imagine, poussant, au-delà de la Méditerranée, jusqu'en Palestine où il aurait formé le peuple des Philistins[7].

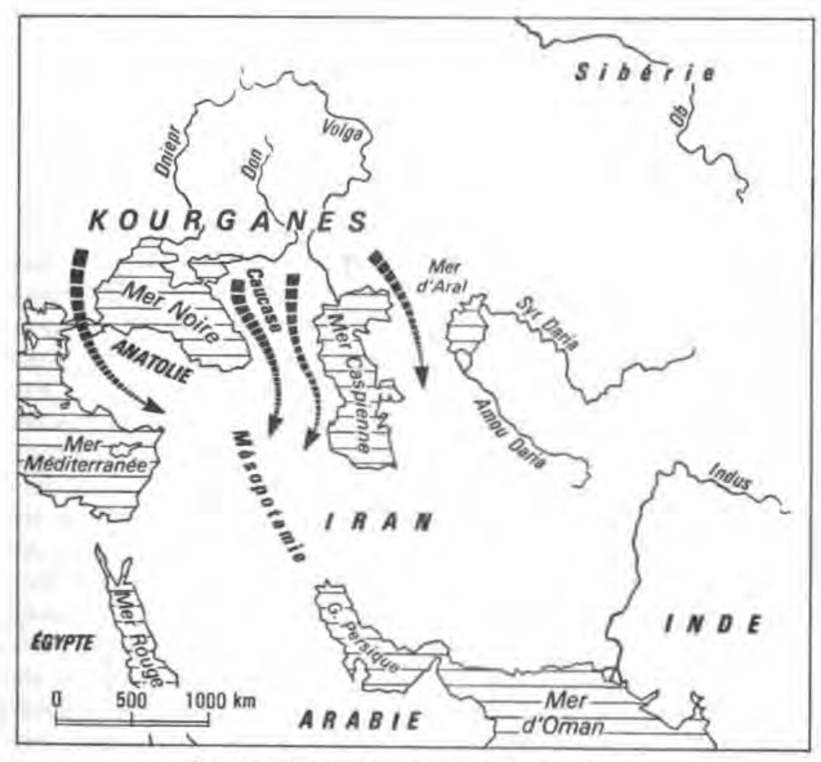
Avec le premier millénaire avant notre ère, nous arrivons à l'âge du fer où toutefois ce métal voisine avec le bronze plus ancien. On distingue une première période dite de Hallstadt, localisée en Autriche centrale, dans la première moitié du millénaire. Linguistiquement, il s'agit de Celtes, d'abord incinérant comme les Villanoviens, mais revenant finalement à l'inhumation directe. Leur centre de gravité ne se déplacera vers l'ouest qu'après − 500, au moment où s'impose la métallurgie du fer. C'est ce qu'on désigne comme la culture de la Tène, du nom d'un lieu-dit situé sur le Plateau Suisse, à proximité du lac de Neuchâtel.

Les Anatoliens.

Les Indo-Européens d'Asie Mineure, ceux qu'on désigne comme les Anatoliens et, notamment les Hittites, y sont probablement parvenus par le Caucase qui, à ses deux extrémités, n'offre guère d'obstacles au passage des peuples en mouvement. Ils y ont établi des contacts certainement très intimes avec les cultures mésopotamiennes, comme en témoigne l'emprunt des graphies cunéiformes. On pourrait penser que les Indo-Iraniens, dont certains au moins ont laissé des traces en Asie Mineure, ont suivi, avec un appréciable retard, la même route

7. Ou faut-il, avec Vladimir GEORGIEV *Introduction to the history of the Indo-European languages*, p. 107, faire remonter beaucoup plus haut dans le temps cette poussée vers le sud et penser aux Pélasges, occupants très anciens de la Grèce, dans lesquels certains veulent voir les représentants d'une ancienne avancée indo-européenne ? Cf., ci-dessous, p. 79.

que les Anatoliens. Mais, d'une part, des considérations linguistiques suggèrent que les ancêtres des Grecs et ceux des Indo-Iraniens sont restés longtemps en contact. Dans un autre sens, lorsqu'on retient qu'aux dates historiques, des peuples Iraniens se sont manifestés dans les steppes, au nord de la Mer noire et très loin en Asie, que très anciennement ils ont été en contact avec les Finno-Ougriens, sans doute aux confins de la Sibérie, on peut se demander si l'on ne doit pas supposer une poussée réalisée par des chemins divers, à l'ouest et à l'est de la Mer noire, voire des deux côtés de la Caspienne, avec convergence finale des courants principaux vers l'Iran et la vallée de l'Indus.



De la région des Kourganes vers le Sud-est

CHAPITRE V

LANGUES ET GROUPES DE LANGUES INDO-EUROPÉENS

On a, jusqu'ici, évité de faire allusion à des langues ou à des peuples peu connus et emprunté des illustrations surtout aux cultures classiques ou occidentales qui ne sont pas étrangères aux lecteurs présumés de cet ouvrage. Tout le monde connaît les Grecs et les Romains, et des termes comme Celtes, Germains ou Slaves ne sont pas sans évoquer quelque chose, même si les idées qu'on se fait des peuples auxquels ils s'appliquaient et continuent de s'appliquer, sont teintées par des préjugés nationalistes apparus le plus souvent sur les bancs de l'école. L'histoire telle qu'on l'a enseignée aux enfants depuis le XIX[e] siècle a visé surtout à renforcer l'unité nationale et, à cette fin, elle a, parmi les données réunies par les chercheurs, fait des choix qui aboutissent nécessairement à une vision partiale et inexacte de la réalité. En France, par exemple, on a choisi, dans l'enseignement destiné au « peuple », d'exalter la Gaule et les Gaulois. Le premier héros national est Vercingétorix, celui qui a tenu tête au Romain César. Dans l'enseignement secondaire, longtemps réservé essentiellement aux « bourgeois », on a conservé une vision plus large de la réalité historique et fait intervenir plus libéralement les origines latines de la culture nationale. Sans pouvoir être nié, le rôle des envahisseurs Germains, dont certains, les Francs, ont, après tout, donné son nom au pays, a été en quelque sorte camouflé : opposer le « sang rouge des Francs » au « sang bleu des Germains » c'est se livrer, inconsciemment peut-être, à une manipulation destinée à faire oublier que, si « germanique » a un sens, cette épithète s'applique aussi bien aux Francs qu'à leurs adversaires alamans de la bataille de Tolbiac (all. *Zülpich*).

On ne saurait poursuivre ici sans préciser ce à quoi on fait allusion lorsqu'on cite telle langue ou tel peuple, soit parce qu'il convient de

rectifier, ou au moins de nuancer, les idées que s'en fait le public, même cultivé, soit parce que seuls les spécialistes en connaissent l'existence.

Dans quel ordre présenter les langues indo-européennes ? Doit-on énumérer toutes celles qui ont une identité en les rangeant de *a* à *z*, ou grouper celles qui, comme les langues romanes, dérivent d'une même langue ancienne, le latin, ou qui, selon toute vraisemblance, ont été, à une certaine époque, des dialectes étroitement apparentés ? Traditionnellement, c'est à ce dernier parti qu'on se range et il n'y a guère de raisons d'agir ici autrement. Mais dans quel ordre devons-nous présenter les différents groupes ainsi délimités ? Allons-nous donner la préférence à une répartition géographique, en retenant les habitats de chaque groupe à la date où ils sont attestés pour la première fois et en procédant d'est en ouest ou d'ouest en est ? Ou donnerons-nous la préférence au temps plutôt qu'à l'espace, à l'histoire plutôt qu'à la géographie ? Et, dans ce cas, choisirons-nous l'histoire de la recherche, en commençant par les langues les plus anciennement utilisées dans la comparaison des langues, ou l'histoire des langues elles-mêmes, en commençant par celles qui, en se séparant plus anciennement du tronc commun, ont assumé plus tôt leur identité ? On peut légitimement hésiter entre ces différents principes d'exposition. Nous prendrons finalement parti pour l'histoire de la recherche, ne serait-ce que parce qu'ainsi nous procédons, en général, du mieux connu au moins bien connu et que nous dégagerons plus tôt les traits qui ont permis d'identifier et de classer les langues découvertes ou déchiffrées tardivement.

La question qui se pose, dès le départ, est de savoir comment nous allons constituer les groupes avec lesquels nous allons opérer : à côté de rapprochements et de groupements qui s'imposent à tout chercheur, il y en a d'autres plus discutables : il n'y a, par exemple, aucun problème en ce qui concerne l'établissement d'un groupe germanique distinct ; à l'intérieur de ce groupe, on n'hésitera pas à opposer un sous-groupe scandinave ou nordique à un autre sous-groupe, dit westique, incluant l'allemand, le néerlandais, l'anglais et le frison. A l'intérieur de ce dernier, on opposera, à l'allemand et au néerlandais, langues de la Teutonia, l'anglais et le frison. Mais si nous considérons les langues, aujourd'hui disparues, que sont le gotique, bien attesté, et le burgonde qui ne l'est guère, on est moins sûr qu'il faille, pour ces dernières, poser un groupe oriental distinct, ou, les rattacher soit au scandinave du fait de certains traits phoniques, soit aux parlers westiques du fait d'identités de vocabulaire. En face de la spécificité du germanique, on trouve le cas du « balto-slave », avec des langues slaves parfaitement caractérisées dont on peut se demander si leurs ressemblances avec les langues dites baltes sont dues, pour l'essentiel,

à une indifférenciation originelle ou à des contacts relativement récents.

Centum et satem.

Si nous voulons bien marquer les rapports et les différences entre les diverses langues, la nécessité d'opérer avec des groupes et des sous-groupes nous incite à commencer par une dichotomie traditionnelle, celle qui distingue, parmi les langues indo-européennes, entre des langues dites *centum* (prononcer *ken'toum*) et les langues dites *satem*. *Centum* est la forme latine du mot *cent* qui, à l'âge classique, ne se prononçait ni à la française, [sētɔm] comme *saint homme*, ni [tšentum], mais avec un [k] initial. *Satem*, ou plus exactement *satəm*, est la forme du même mot en iranien. On voit que le vocable y présente une voyelle sans nasale, mais que, surtout, le [k] initial, qu'on suppose primitif, a évolué vers une sifflante. Ce phénomène n'a rien qui puisse étonner les Français, qui ne distinguent pas, dans la prononciation, le mot *cent* de *sans*. Mais il faut noter que le [k] latin, noté *c*, ne s'est palatalisé, pour donner [tš] en italien, l'interdentale [Þ] en espagnol, [s] en français, que devant des voyelles *i* et *e* articulées en avant dans la bouche, dans la région dite palatale. Dans les langues dites *satem*, ce sont en principe tous les [k] qui se sont palatalisés, quelle qu'ait été la voyelle suivante. Notons que, tout comme dans les langues romanes, le produit de la palatalisation varie d'une langue à une autre. On a *s* en russe comme en iranien, mais [š] en lituanien et une sifflante palatale en sanskrit, le *ch* de *ich* allemand. On notera également que la palatalisation n'affecte pas seulement le [k], mais toutes les articulations occlusives de même profondeur dans la bouche : au *g* du latin *gnō-sco*, *gnōu-ī* « connaître », correspond le *z* du russe *zna-t'* de même sens. C'est là ce qu'on constate également en italien où, en face du [tš] de *cento* « cent », on a l'équivalent sonore [dž] dans *gelare* « geler ». En français et en espagnol, le parallélisme n'a pas été conservé ; on a la sifflante [s] dans *cent*, mais la chuintante [ž] dans *geler ;* l'espagnol présente l'interdentale [Þ] dans *ciento*, mais *h*, c'est-à-dire zéro, dans *helar*.

Il ne faudrait pas croire que la participation à l'évolution *satem* a dû nécessairement impliquer l'appartenance à une unité politique ou ethnique particulière, comme on est tenté de le faire, à tort peut-être, lorsqu'on pose un groupe germanique distinct. Il a pu y avoir parallélisme évolutif chez des gens qui n'étaient plus en contact. Il se trouve toutefois, nous le verrons ci-dessous, que la répartition des langues *satem* sur la carte n'est nullement aléatoire.

On ne dira pas que, parmi les langues indo-européennes, le français est une langue *satem*, parce que l'opposition *centum-satem* se réfère à un processus particulier qui s'est produit à une époque préhistorique et

qui est, dans le temps et dans l'espace, un événement distinct de celui qui caractérise l'apparition des langues romanes. Le premier peut très approximativement être daté du troisième millénaire avant notre ère, le second du IIᵉ siècle après Jésus-Christ. Ceci explique qu'on range également parmi les langues *centum*, le germanique où le *k* ancien est représenté dès les premiers textes par *h*, comme on le voit dans angl. *hundred* « cent ». Que l'on compare également le fr. *corne* et l'équivalent anglais *horn*. Or, ce passage de *k* à *h* n'avait pas commencé au VIᵉ siècle avant notre ère, car le nom du chanvre, qui s'est répandu en Europe vers cette époque et qui a les formes *kannabis*, *cannabis* en grec et en latin, a été emprunté en germanique assez tôt pour que le *k* puisse y évoluer en *h* comme l'attestent les formes modernes de l'angl. *hemp*, all. *Hanf*.

L'évolution en germanique de *k* vers *h* est ponctuée par la forme des noms propres attestée en latin ou en roman. Les Cimbres, connus dès le IIIᵉ siècle avant notre ère, conservent, dans lat. *Cimbri*, leur [k-] initial. Plus tard les *Chatti*, qui deviendront les *Hessois*, ont droit à un [k-] aspiré noté *ch*. Plus tard encore, aux temps mérovingiens, les romanophones vont encore reproduire comme [k], dans *Clovis*, ce qui devait déjà être un [x] (c'est-à-dire le *ach* allemand ou la jota espagnole) et qui, dès les textes allemands du VIIIᵉ siècle, va se trouver noté *h-*, d'où *Hluodowig* (all. *Ludwig*), latinisé en *Ludouicus* et qui aboutit en français à *Louis* (dans la graphie, au *C-* initial près, identique à *Clovis*, écrit en fait *Clouis*). Jusqu'à ce jour, le basque *Koldo* « Louis » conserve le [k-] d'une ancienne forme romane **Klodo-*, avec métathèse du *l*, dans une langue où l'on ne connaît pas de groupe de consonnes à l'initiale du mot.

L'évolution *satem* va de pair avec un autre phénomène. On pose pour l'indo-européen ancien un ordre de labiovélaires, c'est-à-dire des consonnes du type de [k] et [g] accompagnées d'un arrondissement des lèvres qu'on perçoit encore lorsque s'est produite l'explosion de la consonne. C'est pourquoi les Latins notaient la sourde au moyen de deux lettres successives sous la forme de *qu*.

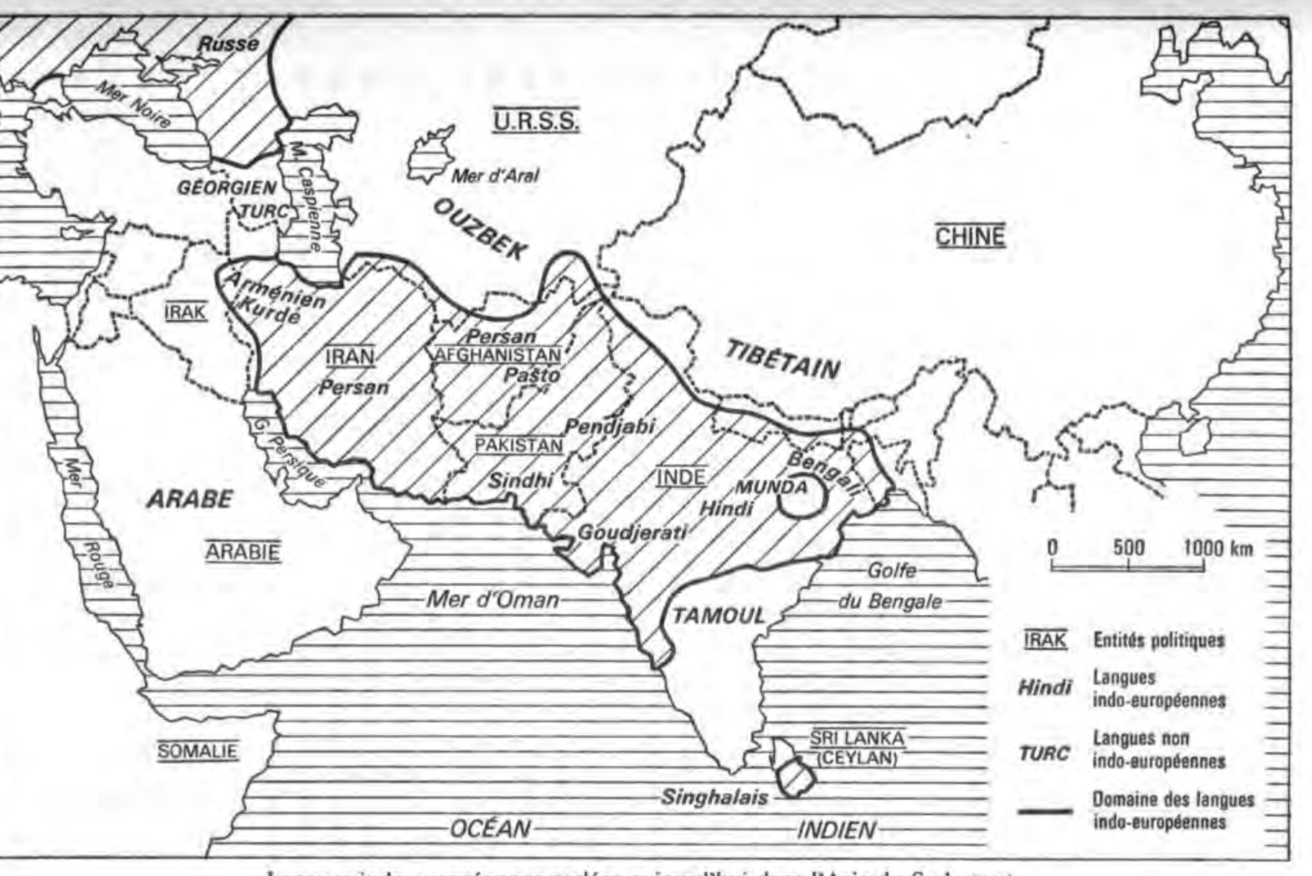
Dans une langue *centum* comme le latin, la labiovélaire, celle qu'on trouve à l'initiale de l'interrogatif *quis*, se maintient. Mais là où [k] se palatalise, *qu* (en notation phonétique [k^w]) se simplifie en [k]. C'est ce qu'on trouve dans les équivalents sanskrit *kas* et russe *k(to)*. C'est également ce qui s'est produit dans les langues romanes où, lorsque *ci-*, dans *ciuitatem* « cité », par exemple, s'est palatalisé pour aboutir au [tš] d'ital. *città*, au [Þ] d'espagnol *ciudad* et au [s] du français *cité*, *quis* a donné *chi* [ki] en italien, *qui(en)* en espagnol et *qui* en français, tous avec [ki] et non [kwi]. Lorsque plus tard, [k] s'est palatalisé en français devant un *a* qui tendait vers [e] (*carum* > *cher*), *qu* y a également perdu son [w], d'où *quand* prononcé [kã], en face de l'italien

quando et l'espagnol *cuando* qui conservent leur [w]. Le germanique a également conservé à date ancienne l'élément labial. C'est le seul élément de l'ancien [kʷ] qui a survécu dans l'anglais britannique *which* [witš], alors que l'usage américain conserve le [h] correspondant au [k] du complexe [kʷ], d'où [hwitš].

La répartition des langues *centum* et des langues *satem*, dans le monde contemporain, se fait de façon assez cohérente : à l'ouest *centum*, à l'est *satem*. Le celtique avec, pour « cent », *cét* [ke:d] en irlandais, *kant* en breton, le germanique *hund-* avec son *h* issu de [k], le lat. avec *centum* [kentum] se situent nettement à l'ouest du domaine. Le grec qui, dans *(he)katón*, a conservé le [k] se trouve aujourd'hui environné de langues *satem*, albanais, serbe, macédonien, bulgare, ces trois dernières slaves. Plus à l'est, le lituanien avec *Šim̃tas*, le russe avec *sto*, l'iranien avec *satəm*, le sanskrit avec *çatam* forment l'ensemble cohérent des langues qui ont palatalisé à date très ancienne les occlusives du type [k]. Rappelons, pour la sonore correspondante, le cas du verbe « connaître » avec le [g] conservé du latin *gnō(scō)*, du grec *(gi)gnṓ(skō)* et l'équivalent régulier [k] du germanique, dans l'anglais écrit *know*, en face des palatales diverses du russe *zna(t')*, du lituanien *žino* « il sait », de l'iranien *žnā(tar-)* « connaisseur » et de son équivalent sanskrit *jnātā-*.

Cette répartition géographique assez cohérente s'est vue battue en brèche par la découverte et l'identification, au début du XXᵉ siècle, dans le Turkestan chinois, d'une langue désignée comme le tokharien. En dépit de cette localisation extrême-orientale, il s'agit d'une langue *centum* où « cent » est *kənte* et où l'équivalent de l'interrogatif *quis* a la forme de *kᵘse* dont le *u* suscrit témoigne du caractère labiovélaire du [k]. Un peu plus tard, en 1917, le déchiffrement du hittite a révélé, en Asie mineure, la présence d'une autre langue *centum* où « cœur » se dit *kartis* avec le [k] du latin *cor*, du grec *kardia*, de l'irlandais *cride* et l'équivalent [h] régulier de l'angl. *heart*, en face de la chuintante et des sifflantes du lituanien *širdis*, du russe *serdce* et de l'arménien *sirt*. L'interrogatif *kwis* conserve en hittite la labiovélaire.

La présence inattendue de langues *centum* à l'est du domaine pourrait s'expliquer, pour le hittite, en faisant valoir que cette langue et les autres langues anatoliennes se sont séparées du tronc commun avant que ne se manifestent les tendances à la palatalisation. Dans le cas du tokharien, pour lequel rien ne semble justifier une séparation aussi ancienne, on peut supposer un déplacement d'ouest en est. On relève le cas amusant d'un mot indo-européen qui n'est attesté qu'aux deux bouts du domaine, en irlandais et en breton à l'ouest, en tokharien à l'est. Il s'agit d'un terme qui désigne la femme ou la jeune fille, qu'on reconstruit comme **kʷljo*, avec ou sans suffixe ajouté : irl. *caile*, et, avec suffixe, *cailín* [kaˈl'i:n], d'où l'anglais *colleen* qui



Langues indo-européennes parlées aujourd'hui dans l'Asie du Sud-ouest

désigne une jeune Irlandaise; en breton, un ancien $*k^{w}l$- a donné régulièrement *pl-*, d'où, avec un autre suffixe, *plac'h*, de même sens; en tokharien, on a $k^{u}lyi$ qui, en dépit des apparences, est le correspondant exact du *caile* irlandais.

LA BRANCHE INDO-IRANIENNE

C'est par cette branche de l'indo-européen que nous commençons, car c'est la découverte du sanskrit, langue de l'Inde, par les Occidentaux qui est à la source de la linguistique comparative. Les peuples qui ont parlé et parlent les langues indo-iraniennes sont souvent désignés comme les Aryens. L'utilisation de ce terme en référence à l'ensemble des « Indo-Européens » est un abus d'autant moins recommandable qu'il a servi aux nationaux-socialistes des années 30 et 40 pour opposer aux Juifs, les non-Juifs. Les Hongrois et les Japonais, alliés des Allemands, ont pu être désignés comme des « Aryens d'honneur ».

Le trait qui oppose le mieux cet ensemble à toutes les autres branches est la confusion en *a*, long ou bref, de deux voyelles, distinguées ailleurs comme *e* et *o*, en grec, en italique et en celtique, comme *e* et *a*, dans les autres langues. Mais on relève beaucoup d'autres points qui indiquent l'existence d'une communauté ancienne groupant Indiens et Iraniens. Cependant, certains faits suggèrent qu'ils se sont très tôt subdivisés entre une branche indienne et une branche iranienne : les chefs du royaume de Mitanni, en Asie mineure, au deuxième millénaire avant notre ère, ont des noms qu'on peut interpréter spécifiquement comme indiens. Cela indiquerait qu'une dissociation a pu s'opérer dès – 2000. Certains chercheurs sont tentés de poser une séparation beaucoup plus ancienne. C'est sans doute à travers l'Iran que les Indiens ont dû gagner le pays dont ils dérivent leur nom, mais pas nécessairement en compagnie des Iraniens qu'ils auraient laissés sur place.

Les Iraniens.

Ceux qu'on nomme les Iraniens parce que c'est en Iran qu'on les connaît à date historique ont occupé ou dominé anciennement une partie importante de l'Asie occidentale jusqu'aux approches de ce qui est aujourd'hui la Sibérie. Ils y sont entrés en contact avec les Finno-Ougriens, représentants d'une autre famille de langue, qui leur ont fait quelques emprunts. La grande langue iranienne de ces régions a été le sogdien. Mais elle n'y était pas la seule. On rattache aux Iraniens ceux que l'antiquité classique a connus sous le nom de Scythes et peut-être,

par leur langue, les Sarmates qui leur ont succédé dans ce qui est aujourd'hui la Russie méridionale. C'est là sans doute que les ancêtres des Slaves ont fait quelques emprunts à l'iranien.

Les langues iraniennes les plus anciennement attestées sont le vieux-perse et la langue de l'Avesta. Le vieux-perse l'est par des inscriptions en graphie d'origine akkadienne, entre le VIIᵉ et le Vᵉ siècle avant notre ère. Il s'agit d'une forme occidentale de l'iranien, la Perse étant proprement la partie de l'Iran située près du Golfe Persique. C'était la langue des souverains dont les armées ont menacé la Grèce. L'Avesta, bible du mazdéisme, est un ensemble de textes religieux, de dates diverses et d'une interprétation linguistique délicate ; certaines parties sont rédigées dans une langue aussi archaïque que les plus anciens textes indiens. Cette langue représente un dialecte plus oriental que le vieux-perse. Aujourd'hui, les langues iraniennes sont, essentiellement, outre le persan, langue officielle de l'Iran, le kurde, parlé en Iran, en Irak et en Turquie, le pašto qui est, en Afghanistan, en compétition avec le persan et déborde sur le Pakistan. Isolé dans le Caucase, se trouve l'ossète qui est peut-être ce qui demeure de la langue des anciens Scythes.

Les Indiens.

Les Indiens, au sens propre du terme, ont dû s'étendre progressivement dans la péninsule dont ils portent le nom aux dépens de peuples de langues dravidiennes non indo-européennes. Ces langues, dont la plus connue est le tamoul, sont conservées jusqu'à ce jour dans le sud-est du sous-continent, avec quelques îlots isolés, vers le nord, jusqu'en territoire iranien.

La forme la plus ancienne des langues indiennes est attestée dans le Rig-Veda, partie la plus archaïque des Vedas, textes sacrés du brahmanisme. La langue de ces textes est connue sous le nom de « sanskrit », en quelque sorte l' « élaborée ». Elle s'est établie comme la langue de culture de l'Inde tout entière, y compris les régions de langues dravidiennes. Les premiers textes sanskrits sont tardifs, mais ils nous offrent des formes linguistiques plus anciennes, transmises oralement, qu'on date volontiers d'avant – 1000. Très tôt ont existé, concurremment avec la langue littéraire, ce qu'on désigne comme les prākrits, c'est-à-dire les parlers quotidiens plus évolués.

Aujourd'hui, il existe plusieurs langues indiennes comme le bengali, le goudjerati, le pendjabi, mais la langue officielle de l'Etat indien, qui s'emploie concurremment avec l'anglais, est le hindi, forme littéraire de l'hindoustani, notée dans la graphie dite *devanagari*, qui sert également pour le sanskrit. Beaucoup de termes y ont été empruntés à cette langue, un peu comme le latin et le grec ont fourni beaucoup de

termes aux langues européennes. Ce qu'on nomme le *devanagari* a coexisté avec d'autres écritures du même type. C'est, plutôt qu'un alphabet proprement dit, un syllabaire où chaque lettre vaut pour une consonne suivie de [a]. Si la voyelle est un [a] long ou d'un autre timbre, on utilise un diacritique postposé ou suscrit. Le sens de la graphie est de gauche à droite comme pour les alphabets occidentaux. La graphie *devanagari* est très typée : l'ensemble des signes d'un même mot présente à son sommet une barre horizontale pratiquement ininterrompue. Une autre forme littéraire de l'hindoustani est l'urdu. Caractérisé par de nombreux emprunts au persan et écrit au moyen de l'alphabet arabe, il sert aux musulmans de l'Inde et du Pakistan.

A signaler, parmi les langues indiennes, le tsigane dont les locuteurs ont dû quitter l'Inde depuis au moins 1 500 ans. On trouve toutefois en Inde des Tsiganes, identifiés comme tels, qui ont le même comportement nomadique que dans le reste du monde.

Ce qui caractérise le sanskrit et jusqu'à un certain point les langues contemporaines de l'Inde, par opposition à l'iranien, c'est son système de consonnes dont la richesse et la symétrie avaient impressionné les premiers comparatistes au point qu'ils l'avaient mis à la base de leurs reconstructions. L'existence, dans ces langues, d'un ordre de rétroflexes, c'est-à-dire de consonnes articulées avec la pointe de la langue incurvée vers l'arrière et placée contre le palais, a vite été interprétée comme un trait spécifiquement indien. Mais les quatre séries d'occlusives, qu'on peut illustrer pour les labiales, par exemple, au moyen de p, p^h, b et b^h, ont été longtemps, et, par certains chercheurs, jusqu'à ce jour, considérées comme un précieux héritage de la « langue mère », c'est-à-dire l'indo-européen commun. En voici un tableau complet, inspiré du sanskrit :

p	t	k'	k	k^w
p^h	t^h	$k^{h\prime}$	k^h	k^{hw}
b	d	g'	g	g^w
b^h	d^h	$g^{h\prime}$	g^h	g^{hw}

On verra ultérieurement quelles réserves il suggère. L'iranien ne connaît pas les rétroflexes et il a réduit à deux (*p, t, ..., b, d,...*) les quatre séries du sanskrit. Contrairement à ce qui se passe en iranien, l'évolution *satem* de l'indien n'a pas abouti à confondre le *s* issu de [k] avec le *s* commun aux langues *centum* et *satem*.

L'ARMÉNIEN

Parlé aujourd'hui encore en Arménie, là où l'Asie mineure confine au Caucase, l'arménien a essaimé du fait des persécutions dont ses



Langues indo-européennes parlées aujourd'hui en Europe

usagers ont été l'objet. On trouve aujourd'hui des communautés arméniennes un peu partout, en URSS, en France, au Liban, aux Etats-Unis et ailleurs. On a cru, pendant quelque temps, que l'arménien était une langue iranienne du fait des nombreux emprunts faits par les Arméniens à leurs voisins du sud-est. La finale *-ian*, si caractéristique des noms arméniens, est en fait d'origine iranienne. En réalité, il s'agit d'une variété extrêmement typée de l'indo-européen, qui a fait une mutation consonantique analogue à celle qu'on retrouve en germanique avec, au départ, aspiration des sourdes (*t'aršamim* « je fane », en face de grec *tarsos* « four à sécher »), dévoisement des sonores (*tiv* « jour » en face de lat. *diēs*) et évolution en sonores simples d'anciennes « sonores aspirées » (*dustr* « fille » en face de grec *thugátēr* où *th-* vient de **dh-*). Mais le parallélisme qui, d'ailleurs, ne se poursuit pas dans le détail, n'implique aucun apparentement particulier entre arménien et germanique. Les évolutions régulières ont souvent abouti à des formes difficilement rapprochables, à première vue, de celle des autres langues de la famille : « deux » se dit *erku*, ce qui correspond point par point au **dwō* initial (lat. *duō*, grec *dúo*) : le *-w-* de **dwō* passe à *-g*^{w}*-* puis à *-g-* qui s'assourdit en *-k-* ; *d*, dans cette position, passe à *r*, la combinaison *-rk-* étant pourvue, pour être prononçable, d'un *e-* prosthétique.

L'ALBANAIS

La langue n'est guère attestée avant le XVIe siècle. Elle fourmille de mots d'emprunt au slave, au turc, au grec et aux parlers romans, si bien qu'il n'y a guère que le dixième du vocabulaire qui soit proprement indigène. L'albanais est une de ces langues qui posent au comparatiste plus de problèmes qu'elle ne contribuent à en résoudre. On a souvent voulu le rattacher à l'illyrien sur la foi de la localisation géographique de l'Albanie. En effet, les Grecs situaient l'Illyrie à peu près où se trouve l'Albanie d'aujourd'hui. Plus tard, les Romains ont placé l'Illyrie plus au nord, dans ce qui est aujourd'hui la Croatie et la Slovénie. Malheureusement les langues parlées autrefois dans ces différentes régions ne nous sont guère connues que par des toponymes qui ne prouvent pas grand-chose. Comme l'étiquette « illyrien » était en fait disponible, on a cherché à la placer sur celles des langues identifiées sur le pourtour de l'Adriatique qui n'étaient pas revendiquées pour d'autres branches de l'indo-européen. Cela a été le cas du vénète de Vénétie et du messapien attesté par quelques inscriptions dans les Pouilles et en Calabre. A l'examen, le vénète s'est révélé proche des autres langues d'Italie et en particulier du latin. Quant au messapien, on manque encore de preuves évidentes de sa parenté avec

l'albanais et l'on voit mal l'avantage qu'on aurait à le baptiser « illyrien ».

On se plaît à rattacher à l'illyrien la langue des Philistins de Palestine. Cette dernière a en grec la forme *Palaistinē*, ce qui rappelle la ville de Palaiste en Epire. *Palaistînos* désigne anciennement le Strymon, fleuve de Thrace. On envisage, vers la fin du second millénaire avant notre ère, un raid venu du nord de l'Europe, dit « des longues épées », mentionné ci-dessus, p. 60, qui, sur son élan, aurait poussé, par terre ou par mer, jusqu'au site de l'actuelle Gaza. Rappelons simplement qu'on ne connaît des Philistins que leur nom et ce que nous en dit la tradition hébraïque.

L'albanais est certainement une langue *satem*, bien que le mot pour « cent » semble y être [k̂ent], avec un [k] palatalisé, qu'on peut supposer emprunté au roman. Le **k* y est attesté comme [s] ou comme [θ], par exemple dans *vis* « endroit » en face de gr. *(w)oikos* « maison » et lat. *uīcus* « quartier », et dans *θɛni* en face de grec *konís* « lente ». Le **g* et le **gh* ont pris les formes sonores correspondant à [s] et [θ], c'est-à-dire [z] et [ð], dans *-zet* « vingt », en face de lat. *(uī)-ginti*, *ðemp*, russe *zub* « dent », en face de gr. *gomphos* « cheville », *erða* « je suis venu », en face de grec *erkhomai* « je viens ». Ces produits ne se confondent pas avec les représentants du **s* primitif qui sont [š] dans les contextes palataux (dans *šī* « pluie » en face de grec *huei* « il pleut », de **suei*) et, ailleurs, tantôt *h*, tantôt son ancien partenaire sonore γ qui a dû apparaître dans des contextes particuliers : on a *hel'k'* « je tire », en face de lat. *sulcus* « sillon », mais *γarpɛr* « serpent », en face de lat. *serpō* « ramper ». Ceci engagerait à rattacher l'albanais plutôt à des langues situées plus à l'est, comme le macédonien ou mieux le thrace, deux langues dont on ne sait pas grand-chose, sinon que la seconde était *satem*.

LA BRANCHE BALTO-SLAVE

Il s'agit là de langues *satem* caractérisées notamment par le fait que les quatre séries d'occlusives établies à partir du sanskrit n'y sont représentées que par deux séries, une série /p/ et une série /b/. C'est ce qu'elles ont en commun avec l'iranien, mais elles s'en distinguent notamment en ce qu'elles ne confondent pas les timbres vocaliques [e] et [a], ni probablement, à date ancienne, [o] et [a].

Le balte.

On doit traiter à part des langues baltiques et des langues slaves. Les premières sont attestées très tardivement, au XVIᵉ siècle, comme

l'albanais. Mais ne s'étant pas, comme ce dernier, trouvées sur des lieux de passage, elles sont plus conservatrices et leur nature de langue indo-européenne est beaucoup plus évidente.

Le vieux-prussien.

Chronologiquement, la première identifiée est le vieux-prussien, disparu aujourd'hui. Il est connu par un catéchisme rédigé par des évangélisateurs germanophones, avec des graphies qu'il faut prononcer comme s'il s'agissait de l'allemand. Soit, par exemple, pour « langue », la forme *insuwis* qu'on doit prononcer [inzuvis], ou pour « poisson » (à l'accusatif pluriel) *suckans* qui doit se lire [zukans].

Le lituanien.

Le lituanien, qui se parle toujours, présente des traits très archaïques. Comme le vieux-prussien, il conserve les finales pleines correspondant aux *-us* du latin ou aux *-os* du grec. On y distingue deux tons, sur les voyelles longues et les diphtongues accentuées, qui correspondent chacun à des allongements de dates différentes. Celui qui est marqué au moyen de l'accent aigu commence sur un sommet mélodique, dans *žinóti* « savoir », par exemple. L'accent circonflexe, sous la forme qu'il a en grec, indique une mélodie qui monte pour atteindre son sommet vers la fin de la voyelle, dans *eĩti* « aller », par exemple. La forme *šim̃tas* pour « cent » illustre plusieurs caractéristiques de la langue : tout d'abord le traitement *satem* du **k* initial ; puis l'existence du ton marqué par le circonflexe sur le groupe *-im̃-* traité comme une diphtongue ; la conservation du *-s* final de nominatif ; mais, ici, le *-as* remplace la finale primitive en **-om*, car cette langue, d'allure générale si archaïque, a innové en confondant, comme les langues romanes, le neutre avec le masculin : d'où le remplacement de la finale neutre *-om* par *-os* qui passe à *-as*.

On a longtemps argué de l'existence de deux tons en lituanien et en grec, pour y voir un trait qui aurait caractérisé l'indo-européen commun. Ce qui est assez vraisemblable, c'est que plusieurs branches, dont le balto-slave, le grec, le germanique, ont fait une évolution parallèle, au cours de laquelle de nouvelles voyelles longues sont apparues dans des conditions qui ont abouti à la différenciation de deux tons : supposons qu'une voyelle longue résulte de la contraction de deux voyelles brèves en contact, dont la seconde est accentuée, soit *-oó-* passant par *-ō̃-*. Si, comme il est probable, l'accent entraîne une montée mélodique, la partie finale de *-ō̃-* sera sur une note plus haute que le début. C'est ce qu'on note, en lituanien comme en grec, au moyen du circonflexe. Mais si une voyelle accentuée est suivie d'une

consonne qui, au cours du temps, s'amuit en allongeant la voyelle, comme l'*s* de l'ancien français *paste* (conservé dans l'emprunt anglais *paste*) a disparu en allongeant le *a* précédent, d'où *pâte*, le point le plus élevé de la courbe mélodique sera naturellement sur le début de la voyelle longue, la partie qui correspond à l'ancienne voyelle. On aura ainsi une descente mélodique, de même qu'on avait, dans le cas précédent, une montée mélodique. C'est une évolution de ce type qu'on doit poser, non seulement pour le balto-slave, mais pour le grec et le germanique. Noter que l'accent, dit « circonflexe », du français *pâte* n'a ni la forme ni la valeur de l'accent circonflexe du lituanien et du grec. Pour des raisons pratiques, on utilise ici, dans la translitération du grec, la forme française du circonflexe.

Le letton et l'anticipation des tons.

Le letton représente, par rapport au lituanien, un stade généralement plus évolué du baltique. Les tons y semblent inversés, le sommet mélodique étant sur le début de la voyelle là où il est sur la fin en lituanien et vice-versa. Cela étonne toujours, car on ne voit pas à première vue comment a pu se produire ce chassé-croisé. En fait, la chose est assez simple si l'on se rappelle qu'une montée mélodique qui contribue à l'identification des mots et des formes est précédée et suivie d'un creux de la courbe pour mieux la mettre en valeur. Il n'est pas rare que les locuteurs anticipent une articulation et tout particulièrement une articulation mélodique. Supposons une telle anticipation graduelle de la mélodie pour l'un et l'autre tons. On part d'une part d'un système où s'opposent un ton *haut à l'initiale* (type *žinóti*) à un ton *bas à l'initiale* (type *eĩti*). Le schéma ci-dessous, où les parties pleines des courbes correspondent à ce qui est perçu parce que coïncidant avec la voyelle, illustre l'évolution du premier ton (type *žinóti*). Le premier temps y correspond à la courbe notée 1. Dans un deuxième temps, le début de la voyelle au lieu de se trouver, pour *žinóti*, sur le point le plus haut de la courbe, va coïncider avec sa descente vers le creux qui suit, ce que représente la courbe 2 du schéma. Pour l'autre ton (type *eĩti*), le début de la voyelle, au lieu de se trouver au point le plus bas de la courbe, va coïncider avec sa montée. Dans ce deuxième temps, un ton *descendant* (dans *žinóti*) va donc s'opposer à un ton *montant* (dans *eĩti*). Dans un troisième temps, le début de la voyelle (dans *žinóti*) va coïncider avec le creux de la courbe. C'est ce que représente la courbe 3 du tableau. Pour l'autre ton (dans *eĩti*), le début de la voyelle va coïncider avec le sommet de la courbe. Il y aura donc opposition d'un ton *bas à l'initiale* à un ton *haut à l'initiale*, comme dans le premier temps, mais les mots où le sommet de la courbe était au début, l'ont désormais à la fin et vice-versa.

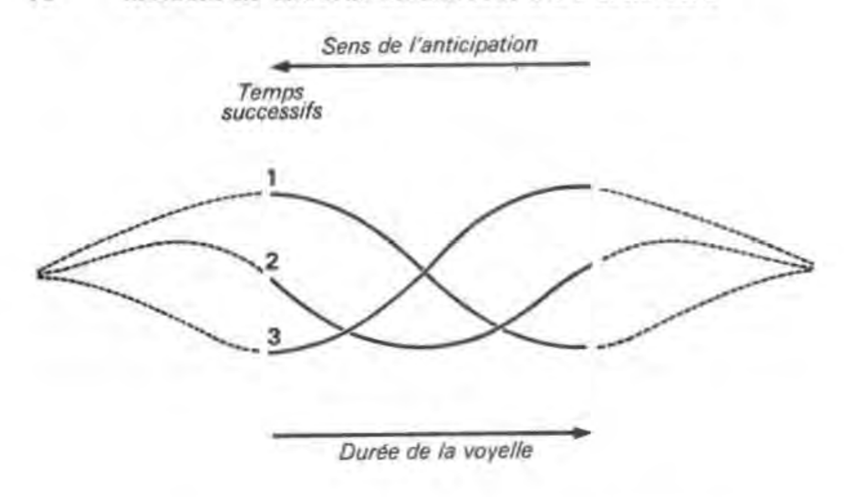


Schéma de l'évolution d'un ton haut à l'initiale
vers un ton bas à l'initiale

Ce sont en général les contacts entre gens de langues différentes qui sont susceptibles de susciter ou de hâter les évolutions linguistiques, et l'on peut conclure de l'archaïsme des langues baltes qu'elles ont été parlées par des populations assez stables situées dans des régions qui ne se trouvaient pas sur le trajet des grands courants migratoires. On verrait donc assez bien les Baltes établis de longue date dans des régions au départ forestières, c'est-à-dire peu favorables au nomadisme des éleveurs de bétail, situées entre les marais du Pripet au sud-est et la Baltique au nord-ouest, celles-mêmes où on les trouve aujourd'hui.

Le slave.

Les langues slaves, aujourd'hui nombreuses et qui occupent un domaine assez vaste, dérivent à une date relativement récente d'une forme de l'indo-européen qui, au début de notre ère, pouvait être parlée dans le nord-ouest de ce qui est aujourd'hui l'Ukraine, donc assez près du site envisagé pour les Baltes, mais dans une région plus exposée aux contacts avec les nomades. Certains emprunts, le mot pour « cent », par exemple, indiquent des contacts avec des Iraniens, probablement des Scythes, établis, nous l'avons vu, au nord de la Mer Noire quelques siècles avant notre ère. Mais il semble que ce soit surtout au contact des Goths, Germains venus de Scandinavie et rayonnant vers le sud-est à partir de la basse Vistule, que les Slaves

ont, dans les premiers siècles après J.-C., reçu l'impulsion culturelle qui sera à la source de leur expansion. Il y a quelques chances pour que des mots comme russe *khleb* « pain » ou *gorod, -grad* « ville » soient des emprunts au gotique *hlaifs* (angl. *loaf*), d'une part, *gards* « maison, lieu d'habitation dans une enceinte » (angl. *yard*), d'autre part. On peut même se demander si l'adverbe *vy* du russe, dont le comportement accentuel comme préverbe est très particulier, n'a pas été influencé par l'équivalent, formel et sémantique, *ūt* (anglais *out*) du gotique. Les progrès techniques résultant de ce contact ont, selon toute vraisemblance, été à la source de l'explosion démographique qu'il faut supposer pour expliquer l'expansion slave ultérieure. Sous la pression des Huns venus d'Asie, les Goths émigrent vers l'Italie, le sud de la Gaule et l'Espagne. Un peu plus tard, les Avars, venus eux aussi d'Asie, semblent avoir entraîné avec eux les Slaves vers l'ouest et le sud. Les Slaves, qui finalement s'imposent, moins peut-être par la violence que par la fécondité, occupent les Balkans où ils se heurtent à Byzance. Vers l'ouest, par le bassin du Danube, ils s'infiltrent jusqu'au Tyrol et occupent une grande partie de ce qui est l'Allemagne d'aujourd'hui. Au-delà de l'Elbe, au niveau de Hanovre, subsistera jusqu'au XVIIe siècle une peuplade slave, les Polabes (de *po-* « sur » et de *-lab-* « l'Elbe », *Albis*, avec anticipation slave de la liquide, d'où **Labis*, en tchèque *Labe*).

C'est l'Empire franc qui le premier va tenter de freiner cette poussée et amorcer ce qui sera la germanisation des pays occupés par les Slaves. On désignera plus tard comme le *Drang nach Osten*, cette poussée vers l'est qui culminera devant Stalingrad au cours de la deuxième guerre mondiale. Vers la fin du IXe siècle, les Hongrois, venus de la Sibérie occidentale, s'établiront dans la plaine du Danube. Ils s'y heurtent vite aux Allemands qui poussent jusqu'à Vienne leur marche vers l'est, fondant l'ancien *Ostmark*, plus tard l'Autriche, *Österreich*. Lorsque les bergers romanophones descendront des Carpathes et peupleront la Moldavie et la Valachie, ils réaliseront la coupure totale entre, d'une part, les Slaves du Sud (Yougoslaves), c'est-à-dire, d'ouest en est, les Slovènes, les Croates, les Serbes, les Macédoniens, et, plus à l'est, les Bulgares slavisés de l'ancienne Mésie, d'autre part, les Slaves de l'est : Russes et Ukrainiens, et les Slaves de l'Ouest : Polonais, Serbes de Lusace, Tchèques et Slovaques.

Au IXe siècle, Cyrille, orginaire de Salonique, et son frère Méthode assument la tâche d'évangéliser les Slaves. Pour traduire les Livres Saints, ils créent successivement une graphie minuscule, dite glagolitique, puis une adaptation des majuscules de l'alphabet grec. Cet alphabet, dit cyrillique, s'imposera finalement chez les Macédoniens, les Serbes, les Bulgares et les Slaves de l'est. La langue employée, dite vieux-slave, permet de toucher les différents slavophones, dont les

parlers sont relativement peu différenciés. Les franges occidentales du domaine, des Polonais aux Croates, évangélisés depuis l'Ouest, adopteront l'alphabet latin.

Le slave est, par rapport au balte, caractérisé par un phénomène qui va se manifester un peu plus tard en français. Il s'agit de l'élimination progressive de tout ce qui, dans la syllabe, se trouve après la voyelle. Le sort d'un mot comme le fr. *fait* illustre bien ce qui s'est passé : sa graphie reproduit sa prononciation en quatre phonèmes /f/, /a/, /i/ et /t/ avec /a/ comme centre de la syllabe ; entre le XIIe et le XVe siècles, le /i/ disparaît en infléchissant la prononciation du /a/ ; le /t/ disparaît, sauf devant la voyelle initiale d'un mot suivant, ce qui permettra plus tard de le rétablir chez ceux qui prononcent [fɛt] dans « C'est un fait ».

Un processus analogue a eu lieu en slave avant le IXe siècle : les consonnes finales tombent ; les diphtongues se réduisent à des voyelles simples ; devant *n* ou *m*, les voyelles se nasalisent avec chute de la consonne nasale : **pontis* « chemin » passe à **pǫ̃ti* et, plus tard, en russe, à *put'* ; les successions *-al-*, *-ar-* passent à *-la-*, *-ra-* ou, en russe, à *-ala-*, *-ara-*, ce qui permet aux syllabes de se terminer en voyelle. On a vu ci-dessus *Albis*, l'Elbe, qui apparaît comme *Labe* en tchèque. Quant à un ancien gotique **gardas*, il va prendre les formes **gradŭ* ou **garadŭ* avant d'aboutir aux formes *gradŭ* du vieux-slave et *gorodŭ* de l'ancienne graphie russe, les deux avec le sens de « ville ». En slave, on va si loin dans ce processus d'élimination des parties finales de la syllabe que les voyelles longues s'abrègent, une partie de leur articulation étant anticipée dans la consonne qui précède : dans une syllabe comme *bī*, [b] va se « mouiller », d'où quelque chose comme [bʲi] ; dans *bū*, l'arrondissement des lèvres de [u] va se reporter sur la consonne, d'où [bʷ] et la voyelle devient la postérieure non arrondie qu'on note *y* dans les translitérations du russe.

Lorsque le processus d'élimination des éléments post-vocaliques est terminé, on va tendre à négliger la prononciation de la voyelle la plus fréquente, ce qui peut se faire sans affecter la compréhension puisque, lorsqu'on entend une consonne sans voyelle suivante, l'auditeur rétablit inconsciemment cette voyelle après la consonne : en français, lorsque *fait* est devenu [fɛ] et *faite* [fɛ-tə], le [ə] de ce dernier a pu disparaître, puisque une prononciation [fɛt] continuait à assurer la distinction entre *faite* et *fait* et qu'encore aujourd'hui beaucoup de Parisiens s'imaginent prononcer le *e* de *faite*. C'est toute l'histoire du « *e* muet ».

En slave, il y avait deux « *e* muets », désignés comme les *jers* (prononcez comme *hier*), un qui reproduisait un *i* bref, l'autre un *u* bref. Ils ont disparu l'un et l'autre, mais le premier a « mouillé » la consonne précédente. Le second lui a donné une coloration [u] qui se perçoit encore nettement dans le cas du *l*, qui est alors « gras »,

comme dans angl. *battle*. De même qu'en français, on conserve certains « *e* muets » pour éviter l'accumulation des consonnes successives, dans *brebis*, par exemple, ou dans *je m' demand'*, on conserve ceux des *jers* qui allègent la prononciation. Ils sont, en russe, représentés par *e*, s'ils dérivent de *ĭ*, par *o*, s'ils dérivent de *ŭ* : un ancien **atĭko-* « petit papa » apparaît en russe comme [at'ets], orthographié *otec* ; l'équivalent russe du vieux-slave *krŭvĭ* (cf. lat. *cruor*) est *krov'* « sang ». Toutefois, le timbre de la voyelle a pu être, dans ce cas, déterminé par la « coloration » de la consonne suivante. Soit, par exemple, le mot qui désigne l'aigle. On peut poser au départ une forme **arĭlŭ*. Le timbre des voyelles va se transférer sur la consonne précédente, les voyelles elles-mêmes restant de timbre neutre et réduites au maximum. Le résultat sera **ar^{j}əl^{u}ə*. La voyelle finale tombe. Il reste donc **ar^{j}əl^{u}*. Celle qui précède va être influencée par la coloration de la consonne qui suit et apparaître, non comme *e*, mais comme *o*, ce qui donne *ar^{j}ol^{u}*, en orthographe *orël*. On trouve des phénomènes analogues en anglais où, par exemple, le mot *silk*, avec un *l* de timbre [u], peut s'entendre comme [suk]. En vieil-anglais, le mot *milk* était noté *mioluc*, le *o* et le *u* visant seulement à marquer le timbre du *l*.

LES AUTRES LANGUES *SATEM*

Certaines langues des Balkans et d'Asie mineure comme le thrace et le phrygien ne sont guère connues que par le nom qui les désignait en grec et par des noms propres. Il est difficile d'en faire état dans la reconstruction de l'indo-européen, non plus que de la langue des Daces et de celle des Gètes parlées autrefois dans la Roumanie actuelle.

LE GREC

La tradition veut que la Grèce ait été occupée par des vagues successives de gens venus du Nord. Nous avons déjà rencontré les Achéens arrivés quelque deux mille ans avant notre ère, et les Doriens apparaissant quelque huit cents ans plus tard. On a parfois envisagé une vague plus ancienne d'occupants indo-européens qui auraient laissé dans la langue des traces caractérisées par une mutation consonantique du type de celle qu'on doit poser pour le germanique et l'arménien. C'est à elle qu'on devrait, par exemple, les finales en *-ntho-* dans *Kórinthos* « Corinthe » par exemple, avec une aspirée au lieu de la non-aspirée primitive des fréquents *-nto-*.

Le grec est une langue *centum*, comme l'attestent le *k* de *hekatón* « cent » et le *g* de *(gi)gnṓskō* « je connais ». Les anciennes labiovélaires du type *k^w ont été, comme il est normal dans une langue *centum*, préservées à date ancienne. La forme *hekatón* présente un trait qui rapproche le grec des langues indo-iraniennes. C'est le changement en *a* des anciennes nasales syllabiques. Le mot qui veut dire « cent » était au départ dérivé du mot qui voulait dire « dix » au moyen du suffixe *-tó-*. Le mot pour « dix » se reconstruit comme **dekm̥ ;* « cent » aurait donc été d'abord **dekm̥tóm* devenu, après chute de la voyelle inaccentuée, **dkm̥tóm*, puis **km̥tóm ; m̥* représente un [m] syllabique, à peu près ce qu'on entend dans une prononciation rapide de *d'après ce que me disent...* [daprɛskm̥diz]. Dans les autres langues, une voyelle s'est développée avant la nasale qui s'est maintenue : on a *e* dans latin *centum*, *a* dans breton *kant*, *u* dans le gotique *hund*, *ə* dans le tokharien *kənte*, *i* dans le lituanien *šim̃tas* qui conserve le *m* primitif assimilé en *n* au *t* suivant dans les autres langues. En grec et en sanskrit, nasale et nasalité ont disparu, d'où, respectivement, *he-kᴀtón* et *çᴀtám*.

Le grec ancien a conservé sous forme d'occlusives les trois séries qu'on reconstruit à coup sûr pour l'indo-européen commun. Soit :

p	t	k
b	d	g
p^h	t^h	k^h

Depuis la fin du monde antique et en grec moderne, les aspirées ont laissé la place à des fricatives : p^h est devenu [f], t^h un [þ] comme le *th* anglais de *thin*, k^h un [x] comme l'*ach-Laut* de l'allemand ou la *jota* espagnole. Dans les emprunts au grec classique, les Français ont [f] pour l'ancien [p^h] dans *phare* par exemple, mais ils ne distinguent pas, dans la prononciation, entre *t* et *th*, entre le *c* (écrit pour *k*) et *ch* (écrit pour *kh*) dans *tome* et *thèse*, *arcadien* et *archéologie*, par exemple. Les prononciations fricatives ont dû apparaître assez tôt en Grèce : on a par exemple, dans des inscriptions de Laconie, des graphies *sio* pour *theos* « dieu ». Mais à Athènes il faut attendre le IIIe siècle après Jésus-Christ pour avoir des prononciations régulièrement fricatives : l'emprunt latin du grec *élephas*, *eléphantos* s'écrit toujours avec *ph*, jamais avec *f*. Cependant l'*olifant* de Roland et *ulbandus* en gotique (où cet emprunt au roman désigne curieusement l'autre animal étrange, le chameau) indiquent que, dans l'usage populaire, on prononçait la fricative. La forme *ol-*, *ul-* manifeste l'évolution régulière en latin d'un *el-* emprunté, comme dans *olīua* « olive », du grec *elaíwā ;* le *-b-* du gotique, qui se prononçait comme un *b* spirant, à l'espagnole, représente l'évolution régulière de *-f-* entre voyelles.

Les données homériques.

Jusqu'à une date très récente, le grec était connu par des textes lapidaires ou manuscrits reproduisant des langues parlées depuis le VI[e] siècle avant notre ère. Certaines formes plus archaïques nous étaient connues par des traditions orales complétées par des données fournies par la métrique. C'est le cas des poèmes homériques, transmis sous une forme qui, en matière de lexique et de grammaire, peut remonter jusqu'au VIII[e] siècle avant notre ère. En matière phonique, les textes nous présentent un usage plus tardif. Mais on peut partiellement reconstituer les formes anciennes en s'appuyant sur la métrique. Tel est le cas du grec *ánax*. On sait qu'en grec comme en latin, les vers sont constitués de successions déterminées de syllabes longues et de syllabes brèves. Une syllabe est longue si elle présente une voyelle longue, comme la première dans lat. *dōnum*, grec *dôron* « don », ou si elle se termine sur une consonne, comme les deux syllabes de lat. *mentum* « menton », grec *anthos* « fleur ». Si, dans le vers, *ánthos* est suivi d'une voyelle (*anthos o...*), le *s* se prononce avec cette voyelle, la syllabe devient *-tho-* et compte donc comme brève. Soit maintenant le grec *ánax* « prince », qu'on trouve dans le nom *Astyanax* (de *ástu* « ville » et *-anax*, donc « prince de la ville »). Si, dans le texte homérique, il est précédé d'un adjectif en *-tos*, on s'attendrait à ce que l'*s* de celui-ci fasse liaison avec le *a-* initial d'*ánax*, la syllabe *-tos*, privée de son *s*, comptant dès lors comme brève. Or, ce n'est pas le cas : *-tos* compte comme longue, comme si *ánax* commençait par une consonne. Effectivement, on sait par d'autres sources que le mot était anciennement *wanax*.

L'alphabet grec.

Le grec classique était universellement écrit au moyen de l'alphabet bien connu, qui est à l'origine de tous les alphabets qui ont suivi. L'alphabet, désigné au moyen de ses deux premières lettres *alpha* et *beta*, est une adaptation heureuse du syllabaire phénicien, inexactement identifié comme un alphabet. En fait, les Phéniciens avaient fort intelligemment constaté que les écritures, probablement mésopotamiennes, dont ils pouvaient s'inspirer, étaient fort mal adaptées à leur langue sémitique. En effet, outre les idéogrammes, elles comportaient des signes correspondant à des syllabes, de telle sorte que les signes pour *ka* et *ku*, par exemple, n'avaient aucun trait commun. Or, en sémitique, ce sont les consonnes qui assurent l'essentiel de l'identité du mot, les voyelles variant au cours de la flexion et de la dérivation : en arabe « il écrivit » se dit *kataba ;* « écrit » se dit *ma-ktūb* avec le

préfixe *ma-* et les mêmes consonnes *k*, *t* et *b*, mais un vocalisme tout à fait différent. Il est paru aux Phéniciens qu'il était préférable d'assurer l'unité du mot en écrivant respectivement *k-t-b-* et *m-k-t-b-* sans noter les voyelles, plutôt que d'employer des signes totalement différents pour les cinq syllabes des deux formes. Ils ont donc établi un syllabaire de 22 signes qui faisait abstraction du timbre des voyelles, *k*, par exemple, valant pour [ka], [ku], [ki], ainsi que pour [k] sans voyelle suivante. Chaque signe portait un nom commençant par la consonne représentée. Le premier de la liste était désigné comme [ʔalef] et valait pour l'occlusion glottale /ʔ/, phonème régulier de la langue. Le numéro six, nommé *wau*, valait pour la consonne /w/, le numéro dix, nommé *yod*, valait pour la consonne /y/. Les Grecs ont reproduit [ʔalef] comme [alpha], dans l'impossibilité où ils étaient de prononcer le [ʔ] qui n'existait pas dans leur langue et, naturellement, le signe correspondant a noté le phonème initial du mot *alpha* désignant la lettre, la voyelle /a/. Le *wau* a été utilisé aussi bien pour le /w/ de *wánax* que pour le /u/ de *phúlax* « gardien » (devenu plus tard, à Athènes, un *u* français). La forme initiale du *wau* s'est ultérieurement différenciée en Y pour la voyelle et F pour la forme non syllabique [w] ; en perdant sa barre verticale, Y a donné V, qui valait en latin aussi bien pour la voyelle de *sus* « porc » que pour la consonne initiale de *uersus* « vers ». Ce n'est qu'au XVIᵉ siècle, en Espagne tout d'abord, qu'on a commencé à distinguer entre *u* et *v*.

Le mycénien.

Les données relatives au grec ont été révolutionnées par une découverte récente : un Britannique, du nom de Ventris, qui, dans l'armée anglaise, avait servi dans le chiffre, a, en collaboration avec un helléniste, Chadwick, appliqué ses connaissances au déchiffrement de ce qu'on appelle le linéaire B. En Crète, dans une région où l'on parlait le grec très anciennement, il se trouvait des inscriptions en deux graphies différentes, désignées arbitrairement comme A et B, qui étaient évidemment des syllabaires. On identifie une graphie comme un syllabaire selon le nombre des caractères. Si une graphie est idéographique, les caractères s'y comptent par milliers. Si une graphie est alphabétique, les caractères s'y comptent par dizaines. S'il s'agit d'un syllabaire, le nombre de caractères distincts se situe entre les deux précédents. Il est donc possible, quand on ne connaît pas une langue, de se prononcer sur le caractère idéogaphique, alphabétique ou syllabique de la graphie utilisée. Ventris a fait l'hypothèse que les graphies en linéaire B correspondaient à du grec et cette hypothèse s'est trouvée vérifiée.

Il est fréquent que, pour réduire le nombre des syllabes notées, on

s'abstienne de tenir compte de certaines distinctions phonologiques, celle par exemple qui existe entre /pa/ et /ba/. En japonais, on utilise dans ce cas des diacritiques : il y a, par exemple, un caractère pour *ha* qui était autrefois /pa/ ; muni d'une sorte de tréma, ce caractère vaut pour *ba*, et avec un petit rond suscrit, il correspond au *pa* qu'on trouve dans les emprunts étrangers comme *pan* « pain ». Dans le syllabaire B, on ne distinguait pas les sourdes des sonores, ce qui laisse, dans bien des cas, la possibilité de différentes lectures. Mais, finalement, on s'y retrouve assez bien et le déchiffrement de Ventris apporte la preuve que le grec du XIVe siècle avant notre ère, dit mycénien, n'avait pas encore, comme le grec classique, réduit les anciennes labiovélaires du type k^w à des labiales (type *p*) ou à des apicales (type *t*), comme on le voit par gr. *leipo* « je laisse » en face du lat. *liquī* « j'ai laissé » ou grec *tís* en face du lat. *quis*. En mycénien, le produit de *$-k^w o$* ne s'identifie pas, comme en grec classique, avec un ancien *-po-*.

On sait que, dans l'antiquité, les Grecs se sont établis en Crète, en Ionie, sur les côtes occidentales de l'Asie mineure, dans l'Italie du Sud, dans ce qu'on appelait la Grande Grèce, en Sicile, et sur la côte provençale et languedocienne, notamment à Marseille, à Nice et à Antibes (*Antípolis* « la ville en face [de Nice] »), à Agde (*Agathḗ* « la Bonne »). Il est inutile de revenir ici sur le rôle décisif qu'a exercé la culture grecque sur le monde méditerranéen et par extension sur la civilisation occidentale, mais aussi, par l'intermédiaire de Byzance, sur l'Europe de l'est. Aujourd'hui, la langue parlée en Grèce est une forme évoluée de la langue commune, la *koïné*, qui s'était établie dans l'antiquité à partir d'Athènes. Jusqu'à ce jour, on hésite entre deux normes, la *katharévousa*, « l'épurée », plus proche de la forme classique, et la *démotique*, ou populaire, plus proche de la réalité parlée.

LES LANGUES ITALIQUES

On a longtemps parlé d'une branche italique de l'indo-européen. Toutefois, certains chercheurs ont émis l'idée que les principales langues indo-européennes parlées anciennement en Italie centrale pouvaient devoir leurs évidentes ressemblances au fait qu'elles s'étaient rencontrées dans la péninsule sans avoir formé, au préalable, un groupement particulier. Nous sommes aujourd'hui trop avertis de l'importance des contacts de langues pour écarter délibérément cette hypothèse. Mais il n'est pas impossible d'expliquer les traits phoniques généraux qui caractérisent ces langues à partir d'un système analogue à celui du grec, avec trois séries de sourdes simples, de sonores et de sourdes aspirées. On pourrait donc supposer, au

troisième millénaire avant notre ère, quelque part en Europe moyenne, un continuum de populations de langue indo-européenne où se serait développé ce système. Une partie de ce continuum aurait migré vers le sud-est et y aurait constitué les Achéens ; le reste, parti en deux vagues vers le sud-ouest, aurait constitué les deux strates de l'italique. La première est peut-être celle qui se trouvait dans les régions marécageuses de la plaine du Pô vers la moitié du second millénaire, donc la population des *terramares*. Ce serait elle qu'on retrouve plus tard dans la péninsule, au-delà des territoires étrusques, et connue sous le nom de Latins. La seconde, porteuse, au début du premier millénaire, de la civilisation dite villanovienne, va graduellement descendre le long des Appennins pour constituer ce qu'on appelle les peuples osco-ombriens. On peut se reporter à la carte, p. 59.

Latin d'une part, osque et ombrien d'autre part, sont des langues *centum*, celles qui, au départ, gardent les labio-vélaires (le type $*k^w$). Mais alors que le latin les conserve, par exemple, dans l'interrogatif *quis*, l'osque et l'ombrien les changent en labiales, d'où *pis* avec la même valeur. La différence entre les deux formes est si frappante qu'on est tenté de lui attribuer une grande importance. Mais l'expérience des diverses langues de la famille montre que le passage de $*k^w$ à [p] n'est nullement exceptionnel, qu'on le constate en roumain, où le latin *aqua* est représenté par *apa*, en Sardaigne, où l'on trouve *battoro* comme produit de lat. *quattuor*, en grec, où le passage de $*\mathit{leik^w\bar{o}}$ à *leípō* « je laisse » a dû se produire après l'époque mycénienne. Nous verrons ci-dessous comment le problème se pose en germanique et en celtique.

L'osque et l'ombrien ont disparu, absorbés par le latin, en laissant peut-être quelques traces dans les parlers de Campanie.

Le latin.

Le latin était, au départ, parlé à Rome et dans sa région, au sud et au sud-est de la ville, dans ce qu'on nomme le Latium. Cette langue était en contact immédiat avec des parlers du type osco-ombrien, le sabin notamment, qui a fourni maintes formes au latin. Il est clair que paysans sabins et paysans romains étaient constamment en contact et le mot qui désigne le « bœuf », *bōs*, a la consonne initiale qu'on attend en osco-ombrien et non en latin, où la forme « régulière » serait $*\mathit{u\bar{o}s}$. Le conflit homonymique avec *uōs* « vous » a pu favoriser l'emprunt. Mais le cas n'est pas isolé. A partir de $*\mathit{w\underset{\circ}{l}k^wos}$ forme reconstruite pour « loup », ce n'est pas *lupus*, avec un *p*, qu'on attendrait en latin, puisque $*k^w$ s'est maintenu dans cette langue.

Le latin apparaît un peu isolé. Il y avait cependant, à dix lieues vers le nord, une ville du nom de Falérie où l'on parlait une langue, le

falisque, assez proche du latin, mais avec quelques traits rappelant l'osco-ombrien, comme le *-f-* intervocalique pour *-b-* latin. La tradition voulait que la ville ait été fondée par les Sicules, établis ensuite dans la Sicile, à laquelle ils ont donné son nom. Si l'on en croit les quelques textes qui nous restent, le sicule ressemblait fort au latin.

L'alphabet latin.

Le latin s'écrit au moyen d'un alphabet dérivé de l'alphabet grec, sous la forme pratiquée par les Doriens de l'Italie méridionale : le H, par exemple, ne désignait pas, comme à Athènes, une voyelle, mais la consonne dite « aspirée ». On attribue à l'influence étrusque la valeur sourde donnée en latin à C au lieu de celle de [g] qui vaut en grec pour la troisième lettre. Il en est résulté la nécessité d'une nouvelle lettre, le G, modification du C. Mais cela n'a pas entraîné l'élimination du K, qui fait double emploi. Le maintien d'un Q, toujours suivi de V, a une justification dans le sens que QVI « qui », monosyllabe, peut ainsi rester distinct de CVI, dissyllabique, qui est le datif du même relatif. Aux 21 lettres traditionnelles, la nécessité de noter les emprunts au grec a fait ajouter Y et Z. On n'a jamais, dans l'antiquité, distingué, dans la graphie, [j] (en français le *y* de *yole*) de [i], [w] de [u]. Lorsque l'alphabet latin a servi à noter d'autres langues, langues romanes résultant de l'évolution du latin, ou langues d'autres groupes, comme le celtique ou le germanique, certaines adaptations ont été nécessaires : on a, par exemple, en germanique, doublé le V pour distinguer le [w] non-syllabique, de la voyelle [u] notée V ; on a eu l'idée de marquer la valeur sifflante de *c* au moyen d'un petit *z* [zɛd] souscrit, d'où le nom espagnol de *zedilla, cedilla* pour désigner cet appendice. Ce n'est qu'au début des temps modernes qu'on s'est avisé de spécialiser chacune des variantes graphiques *i* et *j*, *u* et *v*, pour distinguer entre la voyelle et la consonne. Avant cela, on avait eu recours à un *h* muet pour distinguer entre le *u* de *huile* et le *v* de *vile ;* et, de même, dans *hièble*, une variété de sureau, du latin *ebulum*, pour éviter une prononciation *jèble* qu'aurait suggéré **ièble.*

Les langues romanes.

Comme on le sait, le latin a évolué différemment dans les différentes parties de l'Empire romain où il avait été adopté par les populations locales. Les lettrés continuaient cependant à s'inspirer des normes classiques. Ce n'est qu'au IXᵉ siècle que, dans le nord de la Gaule, on commence à rédiger des textes en langue locale. Les efforts de la royauté franque pour restituer, dans l'usage ecclésiastique, le latin classique, avaient révélé que le bon peuple ne comprenait plus ses

pasteurs, d'où la prescription de dire le sermon en langue vulgaire et la prise de conscience de formes linguistiques distinctes.

On ne saurait ici présenter les différentes langues romanes contemporaines. Il convient simplement de ne pas oublier que leur statut d'aujourd'hui n'affecte pas leur caractère de langues indo-européennes. On rappellera seulement que les langues officielles des cinq États-nations, d'ouest en est, le portugais, l'espagnol, le français, l'italien et le roumain, ne se sont pas, dès l'origine, imposées comme telles, c'est-à-dire comme les seules représentantes valables de la romanité. Il y a eu de multiples tentatives pour élever certains parlers locaux à la dignité de langue littéraire. Rappelons la coexistence prolongée du français et du picard, le développement d'une langue d'oc condamnée par l'hégémonie politique des gens du nord, la présence de trois langues dans la péninsule ibérique, galicien-portugais, castillan et catalan, correspondant chacune à un effort de reconquête, la longue hésitation des Italiens à qui finalement l'autorité de Dante a imposé le florentin. Le roman balkanique subsiste en fait, sous au moins trois formes, le daco-roumain de Roumanie, le mégléno-roumain au nord de Salonique et l'aroumain en Macédoine, sans parler de traces qui remontent vers le nord jusqu'en Istrie. On discute pour savoir si le daco-roumain est bien le produit du latin apporté en Dacie par les légions et préservé dans les Carpathes ou s'il résulte d'un repeuplement par des nomades valaques venus du sud et de l'ouest. Le frioulan et le rhéto-roman, au contact de parlers slaves et germaniques, posent des problèmes particuliers aussi bien en ce qui concerne leur origine que relativement à leur survie. La puissance des États contemporains implique, à plus ou moins longue échéance, la disparition des parlers locaux, surtout lorsqu'ils ne parviennent pas à obtenir un statut politique ou culturel.

LES LANGUES GERMANIQUES

Il est bon de préciser dès l'abord que, dans l'usage linguistique et historique, germanique n'est nullement synonyme d'allemand : l'anglais, le néerlandais et les langues scandinaves sont des langues germaniques de plein droit.

La mutation consonantique.

L'unité des langues germaniques ne peut faire aucun doute. Dès la date relativement tardive où elles sont attestées, c'est-à-dire les premiers siècles de notre ère, elles apparaissent bien typées, essentiellement du fait de la mutation qu'y a subie le système consonantique.

Pour illustrer les différences qui en résultent on rapprochera quelques mots anglais de leurs équivalents grecs et latins : *father* « père » a *f* en face du *p* de *patēr, pater ; heart* « cœur » a *h* correspondant au *k, c* de *kardía, cordis* (génitif) ; *tooth* « dent » a *t-* et *-th* correspondant respectivement au *d-* et au *t-* des génitifs *(o)dóntos, dentis ;* l'anglais a *k-* dans *know* « savoir » en face du *g* du radical *-gnō-* des langues classiques ; noter encore *what* « quoi » avec *wh-* et *-t* correspondant aux *qu-* et *-d* du latin *quod.*

L'allemand a, pour sa part, connu une nouvelle mutation qui lui vaut notamment des sifflantes, *z* et *s*, pour les *t* de l'anglais, dans *Herz (heart), Zahn (tooth)* et *was (what).*

La mutation germanique, qui, dans ses grandes lignes, se retrouve ailleurs, en arménien notamment, comme nous l'avons vu, mérite d'être présentée dans sa genèse et son développement. Tout se passe comme si la glotte tardait à se mettre en mouvement : les cordes vocales, qui constituent la glotte, se rapprochent et entrent en vibration un peu plus tard que prévu : pour [ta] par exemple, ce n'est plus immédiatement après le relâchement du [t], comme pour le français *tas*, mais une fraction de seconde plus tard, l'air ayant le temps de passer avant qu'on ne perçoive la voyelle que caractérise la voix, c'est-à-dire les vibrations de la glotte. C'est ce qui se passe dans la prononciation de l'anglais *tar* qu'on notera [tʰɑ :]. En anglais, toutefois, ce temps est très bref. En danois, en revanche, il s'agit d'une « aspiration » caractérisée et le mot *tager* se notera [thɑʔʁ]. Pour [da], les vibrations sont censées commencer dès le début de la consonne. Ce sont elles qui font l'essentiel de la différence, en français, entre *dard* et *tard ;* dans l'anglais *dart,* elles ont un peu de retard, mais le [d] reste partiellement voisé ; dans le danois *dale,* elles ne débutent qu'avec la voyelle et le [d] s'entend comme un [t] affaibli noté [d̥]. Cependant la distinction reste nette entre *dale* [d̥a : lə] « s'abaisser » et *tale* [tha : lə] « parler », essentiellement du fait de l'aspiration du second.

Tout se passe comme si le danois contemporain était en train de reproduire l'ancienne mutation du germanique commun[1]. Cette mutation a, bien entendu, affecté non seulement *t* et *d*, mais *k* et *g* ainsi que *p* et *b*, ce dernier probablement inexistant en indo-européen commun et n'apparaissant guère que dans des emprunts à d'autres langues. Au cours du temps, les aspirées vont se relâcher, comme elles l'ont fait en grec, d'où le passage à des fricatives : [th] devient ce qu'on note [θ] ou [þ], c'est-à-dire le *th* anglais de *thin ;* de la même façon, **p* passe à [ph], puis à la bilabiale [ɸ] et, ultérieurement, à [f], **k* à [kh] puis à [χ], c'est-à-dire le *ch* de l'allemand *ach* ou la *jota* espagnole ; **kʷ*

1. Pour comprendre ce qui suit, on pourra avoir intérêt à se reporter au schéma de la p. 120.

aboutit de même à [χ^{w}], c'est-à-dire à peu près ce qu'on entend à l'initiale d'esp. *Juan*. A l'initiale [χ] et [χ^{w}] vont s'affaiblir, et ce qu'on entendra sera finalement le frottement de l'air dans la glotte ouverte qu'on appelle l' « aspiration ». Quant aux anciens *b, *d, *g, *g^{w}, après avoir perdu leur voix, ils se renforceront en [p], [t], [k] et [k^{w}].

On ne sait pas ce que pouvaient être exactement, à l'époque de la mutation, les consonnes dites « sonores aspirées » et notées *b^{h}, *d^{h}, *g^{h}, *g^{hw}. Elles apparaissent, à son issue, comme accompagnées de la voix d'un bout à l'autre de leur émission, et, au moins entre voyelles, sans fermeture dans la bouche. On les note souvent au moyen des lettres grecques [β, δ, γ, γ^{w}]. Elles peuvent se durcir en [b, d, g] et [g^{w}] là où elles sont articulées avec une particulière énergie et après une consonne nasale, qui implique une fermeture dans la bouche.

La Loi de Verner.

Avant que l'accent ancien soit remplacé par une mise en valeur automatique de l'initiale du mot, les spirantes sourdes [φ θ χ χ^{w}], issues respectivement de *p, *t, *k, *k^{w}, ainsi que la sifflante *s, se sonorisent en [β δ γ γ^{w}] et [z] lorsqu'elles ne sont *ni à l'initiale ni immédiatement précédées de l'accent.* Les quatre premières se confondent donc avec les produits de *b^{h}, *d^{h}, *g^{h}, *g^{hw}. En conséquence, *km̥tóm « cent » donne, dans un premier temps, [χunθám]. Le [χ-]-initial reste sourd, mais le [-θ-] interne, qui n'est pas précédé, mais suivi de l'accent, se sonorise en [-δ-] qui, en contact avec [-n-], se durcit en [-d-]. Lorsque l'accent se porte sur l'initiale, la finale [-am] disparaît. Le [χ-] initial évolue normalement en [h-]. Le résultat est donc *hund*, attesté tel quel ou dans le composé *hundred*. On doit la formulation de ce faisceau de changements au Danois Karl Verner, d'où le nom de Loi de Verner sous lequel il est connu.

Dans son fonctionnement, le système consonantique qu'on restitue pour le germanique commun et qu'on retrouve effectivement en gotique, ressemble étonnamment à celui de l'espagnol contemporain. Dans le tableau qui suit on trouvera, dans chaque case, un mot gotique et un mot espagnol dans leurs graphies respectives, mais présentant, à l'initiale, le même produit phonique :

[f] *fadar* *fuego*	[θ] *þaurp* *cerro*	[χ] [h] *heþjo* *jefe* *	[χ^{w}] [hw] *hweila* *juego* *
[β] [b] *badi* *beso*	[δ] [d] *daigs* *dedo*	[γ] [g] *galwo* *gana*	[w] [γ^{w}] [g^{w}] *warmjan* *hueso*
[p] *paida* *padre*	[t] *taihum* *tal*	[k] *kara* *casa*	[k^{w}] *qino* [kw] *cuanto*

* Tout particulièrement dans certains usages d'Amérique.

Les produits de -ww-

A l'aube des temps historiques, on peut imaginer les Germains dans ce qui est aujourd'hui le Danemark, le sud de la Suède et les parties de l'Allemagne du Nord à proximité immédiate. Dans la péninsule scandinave se trouvent ceux qui durcissent en *-ggw-* les *-ww-* de la langue. Pour une notion pour laquelle le français n'avait pas de mot jusqu'à la création, récente, de l'adjectif *fiable*, on va trouver, chez ceux qui ont longtemps habité la Suède méridionale, les formes *triggws* du gotique, *tryggr* du vieil-islandais, respectivement *trygg* et *tryg* en suédois et en danois d'aujourd'hui. Chez les descendants de ceux qui occupaient *alors* le Danemark, le Schleswig, le Holstein et le Mecklembourg, les formes attestées sont sans *g(g)* : *triuwe* en vieux-frison, par exemple, *true* en anglais, *treu* en allemand. Le féminin correspondant, *triggwa* en gotique, *triuwa*, puis *Treue* en allemand, a été emprunté par les langues romanes au sens de « trêve ». L'italien et l'espagnol qui tiennent le mot du gotique (Ostrogoths d'Italie, Wisigoths d'Espagne) ont la forme *tregua* qui conserve le *g*. Le français a reçu des Francs la forme sans *g*, *triuwa*. La diphtongue *iu* de l'ancien germanique était accentuée sur le *i*, exactement comme la diphtongue *ie* du vieux-français, dans les deux langues les timbres de *u* et de *e* étant à peine perçus. *Triuwa* a donc été reproduit comme *triewa*, qui a évolué régulièrement en français jusqu'à *trève*.

Les Germains de l'est.

Parmi les Germains établis en Suède méridionale, ceux qui ont changé *-ww-* en *-ggw-*, il en est qui partent, à travers la Baltique, vers le sud et le sud-est. Certains d'entre eux, les Goths, paraissent provenir de l'île de Gotland, au centre de la Baltique. Nous les avons déjà trouvés à l'embouchure de la Vistule, puis dans l'Ukraine d'aujourd'hui, où ils soumettent ceux qui seront les Slaves ; enfin, après avoir traversé l'Europe moyenne, ils s'établissent en Italie, en Gaule méridionale et en Espagne. De l'île de Bornholm, anciennement *Burgundrholm*, aujourd'hui politiquement danoise, mais appartenant physiquement à la Suède, partent les Burgondes. On les retrouvera sur le Rhin moyen où ils fonderont un royaume, puis, plus loin au sud-ouest, de la Bourgogne à la Savoie, où ils se maintiennent quelque temps avant d'être absorbés par les Francs. Goths et Burgondes parlaient une variété de langue, désignée comme le germanique de l'est, qui a disparu aujourd'hui. On la connaît essentiellement par la traduction du Nouveau Testament réalisée, au IVᵉ siècle, par Wulfila, d'une famille de Cappadoce enlevée par les Goths au cours d'une

expédition. Un voyageur flamand a, au XVIe siècle, relevé quelques mots d'un dialecte gotique encore parlé en Crimée à cette époque.

Les Germains du nord.

D'autres, parmi ceux qui sont établis en Suède, restent sur place, ou, traversant les détroits, vont gagner le Danemark. Les linguistes désignent leur langue comme le germanique du nord et ils sont ceux qu'on désignera plus tard comme les Scandinaves. Dans ce terme, on retrouve, légèrement déformé, le nom de la Scanie, la plus méridionale des provinces suédoises. Leur langue est attestée, depuis le IIIe siècle, sous la forme d'inscriptions runiques. La plus ancienne, gravée sur une corne d'or aujourd'hui disparue, peut être reproduite comme *ik hlewagastiz holtingaz horna tawido* et comprise comme « moi Hlewagastiz, fils de Holt, ai fait la corne ». *Hlewagastiz* est formé de *hlewa-*, qui correspond en grec *kléo-s* (< *kléwo-s*) « célèbre », et de *-gastiz* (angl. *guest*, all. *Gast*), qui est le latin *hostis*, mais avec le sens de « hôte » et non d'ennemi.

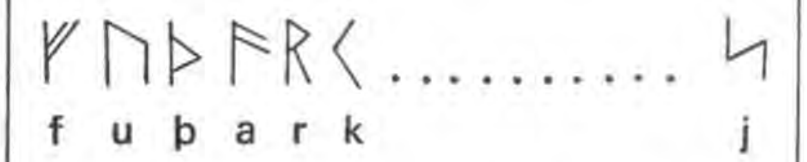
Les runes dérivent de toute évidence d'un des nombreux alphabets apparus en Italie et qui remontent en dernière analyse à celui des Doriens de la Grande Grèce. Les caractères runiques se distinguent de leurs modèles méridionaux par le fait qu'ils évitent toute barre horizontale. Le T, par exemple, conserve sa hampe verticale, mais cette hampe est surmontée de deux barres obliques descendant, à droite et à gauche : ↑. On explique ce trait du fait qu'avant d'être taillées dans la pierre, les runes étaient gravées sur des tablettes de bois. Les hampes verticales, à travers les fibres, étaient très visibles, mais les barres horizontales se confondaient avec elles. Ces tablettes étaient en bois de hêtre (all. *Buche*), d'où le mot *Buchstabe* qui, aujourd'hui encore, désigne la lettre en allemand. En vieil-anglais, *bōc* désigne aussi bien le livre que le hêtre ; en gotique, *boka* est la lettre et son pluriel *bokos* est « le livre ».

Un demi-millénaire plus tard, les Germains du nord, connus alors comme les Vikings, après avoir occupé l'essentiel de la Norvège, iront piller les côtes des Iles britanniques et de la Manche, et s'établiront de façon permanente en Islande, en Angleterre et en Normandie, pour pousser ultérieurement jusqu'en Sicile. A l'est, sur la route de Byzance, ils fonderont l'État russe qui recevra leur nom, *Rusĭ :* en finnois, *Ruotsi* veut dire « Suédois ».

Les Germains de l'Ouest.

Ceux des Germains qui n'ont pas durci *-ww-*, établis tout d'abord au Danemark et, en Allemagne, sur la côte entre la Weser et l'Oder, vont

LES RUNES



Voici les six premiers et le douzième caractères de l'alphabet runique original, dit *futhark*. Lorsqu'on se rappelle que les traits horizontaux étaient exclus, parce que, sur les tablettes, ils se confondaient avec les veines du bois, les runes 1, 4, 5 et 6 s'expliquent aisément à partir des lettres F, A, R et C de l'alphabet latin. La rune 2 est un V renversé où l'on a favorisé les verticales. La rune 3 ne correspond pas à un phonème du latin. On pense naturellement au thêta grec, θ, qui a dû commencer à se prononcer comme le *th* de l'anglais *thin*, au lieu de l'aspirée primitive, vers le début de notre ère.

Chacune des runes avait son nom, et la valeur phonique de chaque rune était soumise aux vicissitudes de l'évolution phonique de l'initiale de ce nom. La rune 4, par exemple, portait le nom des divinités connues plus tard comme les Ases. La forme ancienne du singulier de ce terme devait être [ansaz]. Dans les dialectes germaniques du nord [an] est passé à [ã] devant [-s-], d'où [ãsaz], et, pendant assez longtemps, la rune 4 a eu valeur de [ã] ([a] nasal noté *an* en français). Se posait alors le problème de la notation du [a] non nasal. C'est là qu'intervient la douzième rune. Elle avait reçu le nom de l'année [ja : ra] (< **yēro-* ; cf. ci-dessus p. 53 note 4 à propos de Héra), d'où l'angl. *year* et l'all. *Jahr*. En Scandinavie, le [j] initial (*y*- français) est tombé, le mot a pris la forme [a : ra] (danois et suédois *år*) et la rune ainsi désignée à été utilisée pour noter le [a] ordinaire.

au cours des siècles, essaimer vers le sud et l'ouest. Les linguistes désignent en conséquence leur langue comme le germanique de l'ouest ou westique. Nous avons déjà suivi les Cimbres et les Teutons dans leur dramatique migration du IIᵉ siècle avant notre ère. Dès cette époque, on peut supposer les Germains au contact des Celtes aux abords de la Thuringe. Mais, très vite, l'expansion se poursuit. César arrête un peuple germanique, les Suèves, qui, guidé par son chef, Arioviste, est sur le point d'envahir la Gaule. Les Suèves donneront plus tard leur nom à la Souabe, au sud-ouest de l'Allemagne. A la chute de l'empire romain, on les retrouvera dans la péninsule ibérique, jusqu'au Portugal. Au nord-ouest, les Germains occupent très tôt la côte de la mer du nord jusqu'à l'embouchure de la Meuse où on les identifie comme les Bataves.

Les Anglo-Frisons.

Les peuples de la mer du nord seront désignés plus tard comme les Anglo-Frisons. Ce sont eux qui, au Vᵉ siècle, occuperont le sud-est de la Grande-Bretagne. Linguistiquement, on les reconnaît à ce qu'ils palatalisent les [k] en [ṯ] puis [tš] : la désignation de l'église au moyen d'un mot grec, *kūri(a)kon* « (la maison) du Seigneur », aboutira en anglais à *church*, en face du danois *kirke* qui est à l'origine de l'anglais *Kirk* qui désigne l'Eglise d'Ecosse ; l'all. *Kirche* a traité à sa façon le second *k*. Le latin *caupō* « cabaretier » est à la source de l'all. *kaufen* « vendre », mais aussi de l'angl. *cheap*, dans le nom du quartier de Londres appelé *Cheapside* « le côté du marché » et, naturellement, dans *cheap* « (bon) marché ». On a bien sûr le même traitement dans un mot non emprunté comme *chin* « menton » en face d'all. *Kinn* de même sens, et danois *kind*, got. *kinnus* du sens de « joue », comme le correspondant latin *gena*.

Les Francs.

Avant même que ces peuples de la mer partent coloniser la Grande-Bretagne, se forme, sur leurs arrières et plus au sud, la confédération des Francs. A l'orée du Vᵉ siècle, au moment où vont céder les frontières de l'Empire romain, les Francs sont sur le Rhin, là où il débouche sur la plaine du nord. Sans doute sont-ils, de longue date, présents à titre de colons sur la rive gauche du fleuve. Dès le début du Vᵉ siècle, ils s'avancent en force vers l'ouest et s'établissent solidement dans les territoires qu'ils occupent. Leur première capitale est Tournai, puis Soissons et finalement Paris. Le coup de maître de leur chef, Clovis, est la conversion au christianisme romain de ce peuple, nominalement chrétien, mais sectateur de l'hérésie arienne. Ceci, qui

va faciliter la pénétration mutuelle des Gallo-Romains et des envahisseurs, contraste avec ce qui se passe, vers la même époque, en Espagne où, pendant des siècles, des Wisigoths, nouveaux venus, vont vivre côte à côte avec les premiers occupants, chaque composante de la population conservant ses coutumes propres. Quant aux Francs, leur amalgame avec les Gallo-Romains, joint à leurs talents militaires, leur assure d'emblée une supériorité sur les autres Germains qui ont pénétré sur le territoire de l'Empire. A l'ouest, ils défont les Wisigoths à Vouillé, près de Poitiers. Au sud-est, ils soumettent les Burgondes. Ils battent, non loin du Bonn d'aujourd'hui, à la bataille dite de Tolbiac, les Alamans, nouveaux venus sur la scène. Les Alamans ont pris le relais des Suèves sur le Rhin, haut et moyen, et, avant d'être finalement intégrés à l'empire franc, ils vont occuper ce qui est aujourd'hui l'Alsace et la Suisse. Vers l'est, les Francs étendent également leur domaine dans tout ce qui est l'Allemagne moyenne et notamment la Franconie, au-delà de Francfort (le « Gué des Francs » sur le Main). Plus tard, sous de nouveaux chefs, les Carolingiens, ils refouleront les Arabes jusqu'en Espagne et amorceront la reconquête catalane. Ils pénétreront en Italie où ils rencontreront les Lombards, d'autres Germains de l'ouest qui y avaient succédé aux Ostrogoths, et Charlemagne (*Karl der Grosse*) ira à Rome se faire couronner Empereur d'Occident. Au nord-est, ils contiendront d'autres Germains, les Saxons, ce qui restait, sur le continent, de ceux qui s'étaient joints aux peuples de la mer pour aller occuper l'Angleterre.

Le haut-allemand.

Linguistiquement, une nouvelle coupure s'est établie parmi les Germains d'Allemagne, du fait de ce qu'on désigne comme la seconde mutation consonantique. C'est celle qui donne, pour la charrue, l'allemand *Pflug* en face d'anglais *plough*, danois *plov*, *zahlen* pour angl. *tell*, dan. *tale*, *zwei* pour angl. *two*, dan. *to*. Les Francs, tout comme leurs adversaires saxons et les émigrés de Grande-Bretagne, conservent l'ancien consonantisme. Les emprunts que fera le français à leur langue en témoignent : le « jugement de Dieu », l'*ordalie*, conserve le *-d-*, comme l'anglais *ordǣl*, aujourd'hui *ordeal*, en regard du *-t-* de l'équivalent allemand *Urteil* « jugement ». Les Alamans de Suisse vont le plus loin dans la mutation : ils font « éclater » non seulement le *p* en [pf] et le *t* en [ts] (noté *z*), mais le *k* en *kch* (avec le *ch* de l'all. *ach*) que les Parisiens entendent comme *kr*. Les Lombards participent également à la mutation, comme en témoigne, par exemple, l'ital. *zanna* « dent, défense » avec *z* comme l'all. *Zahn*. On opposera désormais, d'une part, le bas-allemand, celui de la plaine du nord, dont le néerlandais contemporain, des Pays-Bas et de Belgique,

est une variété occidentale, mais qui s'étend jusqu'à Berlin et au-delà, et, d'autre part, le haut-allemand, celui des reliefs accusés de l'Allemagne moyenne et du sud, de l'Autriche et de la Suisse. L'allemand officiel d'aujourd'hui est une création un peu hybride sur la base des usages de l'Allemagne moyenne. Utilisée d'abord comme langue de chancellerie, elle s'est imposée lorsque Luther l'a adoptée pour sa traduction de la Bible.

Le bilinguisme de l'Empire franc.

La symbiose des Francs et des Gallo-Romains du nord a déterminé une longue période de bilinguisme qui a laissé des traces aussi bien en français qu'en allemand : le latin *altus* « haut » avait abouti en Gaule à **aut-*; la forme germanique de même sens était *hauh-*, le *h-* initial de cette dernière a été transféré à **aut*, d'où *haut*; le Midi a, sans *h*, *Auterive*; en Savoie, en dépit de la graphie, on dit *d'Hauteville* et non *de Hauteville*. En anglais, le mot *sake*, par ex. dans *for my sake* « pour l'amour de moi », n'a jamais eu que le sens de « cause » (pour laquelle on prend parti) et ne se confond nullement avec *thing* « chose ». L'équivalent allemand *Sache* est au moins aussi courant que *Ding* avec le sens de « chose », probablement sur le modèle du français où *cause* et *chose* n'ont longtemps été que des variantes du même mot. Le français et l'allemand ont en commun l'emploi du nominatif du mot pour « homme » avec une valeur d'indéfini : fr. *on* de lat. *homo*, all. *man* en face de *Mann* « homme ». L'anglais ne présente rien de semblable.

En ce qui concerne les Germains d'Angleterre, désignés globalement comme les Anglo-Saxons, leur langue a été influencée, plus profondément qu'on l'a longtemps cru, par le celtique qu'ils ont trouvé dans l'île. Puis une longue période de bilinguisme a suivi la conquête normande. Un dictionnaire de l'anglais contemporain contient plus de mots d'origine française ou latine que de termes qui dérivent du germanique, même si ceux-ci sont plus fréquents dans les textes.

LES LANGUES CELTIQUES

On a déjà eu l'occasion, ci-dessus, de mentionner les Celtes à plusieurs reprises, ainsi que le trait le plus frappant de leurs langues, la chute du *p* à l'initiale et entre voyelles. Comme le *p* était d'une grande fréquence en indo-européen commun, sa disparition ne passe guère inaperçue. Dans les langues de l'ouest, le gué, réalité quotidienne avant que ne soient construits des ponts, était désigné au moyen d'un

terme **pr̥to-* dérivé de l'adverbe *per* « à travers ». En germanique, ceci aboutit à *ford* qui est la forme de l'anglais contemporain, attesté, par exemple, dans *Oxford*, le « gué des bœufs » et, nous l'avons vu, dans *Francfort* avec le *-t* de l'all. *Frankfurt*. En celtique ceci donne, selon les normes de l'évolution, **rito-*, avec, pour ce nom neutre, un pluriel **ritā* qui est l'ancien nom de Limoges, sur la Vienne, là où elle devait offrir plusieurs traversées possibles. On retrouve le terme dans *Chambord*, de **camborito-*, le gué sur la courbe du fleuve, comme *Cambridge* est le pont dans un site analogue. Rappelons le cas d'Arles, *Arelate*, à rapprocher des formes grecques *pará* et *platús*. La grande forêt de l'Allemagne moyenne était dite *Hercynienne*, avec un *h-* initial, dernière trace de l'ancien *p-* de **perkus* « chêne » ; avec un *f-* initial régulier, le gotique *fairguni* désigne la montagne, « montagne » et « forêt » étant souvent confondues. Rappelons également **wer-*, dérivé de **uper* « dessus » avec chute du *-p-*, en face de **wo-*, de **upo* « dessous » avec la même chute, l'un et l'autre correspondant au grec *hupér* et *hupó* qui ont un *h-* qui n'est pas nécessairement étymologique.

C'est essentiellement dans l'Allemagne d'aujourd'hui — à l'exception des zones les plus septentrionales —, en Autriche ainsi qu'en Bohême, qu'il faut placer les Celtes à l'aube de l'histoire. C'est de là qu'ils vont rayonner dès le début du premier millénaire avant notre ère, notamment vers l'ouest dans ce qui deviendra la Grande-Bretagne, la Gaule et l'Espagne. Mais aussi vers le sud et l'est : on les retrouve en Italie du nord et, après une équipée en Grèce, en Asie mineure où ils sont connus sous le nom de Galates.

Comme il n'y a, à date ancienne, pas de *b* dans la langue, ***g^w va généralement passer à *b*, comme dans la désignation du bœuf, irl. *bó*, breton *buoc'h*. Lorsque *p* disparaît, la place est libre pour une évolution parallèle de ***k^w, **kw* vers *p*, et c'est effectivement ce qu'on constate, mais dans une partie seulement du domaine celtique : l'équivalent de lat. *quattuor* « quatre » est en vieux-gallois *petguar* avec *p-*, mais en irlandais *cethir* avec *c-* (=[k]) ; le cheval est *epo-* dans les anthroponymes gaulois, mais *ech* (de *ek-*) en irlandais. Les linguistes distinguent ainsi entre un celtique à *q*, ou gaélique, et un celtique à *p*, ou brittonique, le premier situé aujourd'hui en Irlande et en Ecosse, le second au Pays de Galles et en Armorique. Comme le gaulois est considéré « langue à *p* » et que le celtique d'Espagne, mal connu, semble avoir été « à *q* », on estime souvent qu'une première vague, qui avait conservé ***k^w, a poussé jusqu'à l'extrême ouest représenté par l'Irlande et l'Espagne, alors qu'une deuxième vague, « à *p* », s'est établie, moins loin, en Gaule et en Grande-Bretagne. Mais, nous l'avons vu, il n'est pas du tout certain que le gaulois ait été uniformément « à *p* », et on peut supposer que [k^w] et [p] ont pu,

pendant longtemps, être des variantes plus ou moins individuelles du même phonème.

Les labiovélaires.

Il faut se représenter les labiovélaires du type [k^w], [g^w] comme des formes très instables susceptibles à la moindre provocation de virer à [p] et à [b]. Le cas de ce qui est, dans maintes langues, la désignation de l'éponge est, à cet égard, symptomatique. On doit, semble-t-il, partir d'un original *$g^{hw}ong^{hw}o$-, forme expressive à consonne redoublée désignant un amas de matière molle, cédant à la pression, comme le sol du marais, un champignon, ou des muqueuses, les lèvres par exemple[2]. Cette forme pourra recevoir un *s-* préfixé, fréquent en indo-européen, comme dans **stegō* à côté de **tegō* « couvrir », sans qu'on puisse lui attribuer une valeur déterminée. On part donc de *$g^{hw}ong^{hw}o$- ou de *s-$g^{hw}ong^{hw}o$-, mais on trouve aussi une forme « féminine » en *-ā* final au lieu de *-o*. Il va se produire une dissimilation d'un des deux *g^{hw} par l'autre, c'est-à-dire que l'un va se différencier en *b^h. Si c'est le premier, on obtiendra *$sb^hong^{hw}o$ qui apparaît en grec comme *sphóggos* (c'est-à-dire *sphóngos*) aussi attesté comme *spóggos*, *spoggiá*, formes qui sont à la source des formes latines correspondantes, d'où le français *éponge* et l'anglais *sponge*. Si c'est le second g^{hw} qui cède et passe à *b^h, on aura *$g^{hw}ombho$-, *$g^{hw}ombha$, celui-ci donnant en russe *guba* « éponge », mais aussi « lèvre ». Avec *s-*, on aura s-$g^{hw}ombho$- qui donne en germanique **swamb-*, d'où l'allemand *Schwamm* « éponge », mais aussi « champignon ». On trouve des formes parallèles, avec *-p-* au lieu de *-b-*, le danois *svamp* « éponge » et « champignon », l'anglais *swamp* « marais ». D'une forme à voyelle réduite en première syllabe *s-$g^{hw}\c{m}b^h$- on attend **swumb-* puis **sumb* et, avec *p* au lieu de *b*, **sump*, d'où l'allemand *Sumpf* « marais ». Avec un traitement particulier de l'initiale, on a, en grec, *somphós* « spongieux ».

Au lieu d'une dissimilation, on peut avoir au contraire assimilation d'une labiale initiale à une labiovélaire qui suit à distance : le latin *quercus* « chêne » remonte à un **querquus* qui dérive du *$perk^wus$ que nous avons rencontré ci-dessus à propos de la forêt Hercynienne. De même dans le latin *quinque* « cinq » en face du grec *pénte* (= *$penk^we$) : l'irlandais *cóic*, de même sens, remonte également à une forme à deux *k^w.

2. On revient longuement sur ce prototype dans La phonétique des mots expressifs ; le cas d'*éponge*, à paraître dans *Glossologia*. Athènes, 3 (1986). Le prototype y est reconstruit comme un redoublement du type *traintrain*, *gnangnan*, la consonne initiale de chacune des deux syllabes étant un g^{hw} prénasalisé ; cf., ci-dessous, p. 169-173.

La lénition.

Les langues celtiques, dans leur ensemble, ont été affectées, dès avant notre ère, par ce qu'on appelle la lénition. Il s'agit de l'affaiblissement des consonnes lorsqu'elles se trouvent entre deux voyelles. Le processus peut être illustré au moyen d'exemples empruntés aux langues romanes : soit lat. *sonator* « celui qui sonne » ; le *t* est intact dans l'italien *sonatore,* mais il a été affaibli dans l'espagnol *sonador* en un *d* d'articulation très faible (le *th* d'anglais *this* en moins énergique) ; dans le français *sonneur,* l'affaiblissement a abouti à la disparition de la consonne.

La lénition se produit dans les langues qui présentent beaucoup de géminées, c'est-à-dire des groupes de deux consonnes identiques entre voyelles. Beaucoup de Français prononcent un *l* géminé dans *illusion* et un *m* géminé dans *sommet.* Il peut se produire que, du fait de l'évolution de la langue, le nombre des géminées s'accroisse considérablement dans une langue, au point que la succession *-atta-* y soit du même ordre de fréquence que *-ata-,* ou, en tout cas, qu'elle y apparaisse beaucoup plus souvent qu'un groupe de consonnes quelconques, celui de *-asta-* par exemple. Comme ce qui est fréquent est peu informatif, les gens vont tendre inconsciemment à réduire l'énergie utilisée pour la production des géminées devenues très fréquentes : *-atta-* tendra donc vers *-ata-.* Mais comme il faut continuer à se comprendre et, en français, par exemple, ne pas confondre *la dent* /ladã/ avec *là-dedans* /laddã/ ou *il a pris* /ilapri/ avec *il l'a pris* /illapri/, la géminée ne pourra s'affaiblir que si la simple placée dans la même position, donc entre voyelles, s'affaiblit en même temps ; *-atta-* ne deviendra donc *-ata-* que lorsque l'ancien *-ata-* sera devenu, selon les langues *-ada-,* ou *-atha-,* avec un *th* analogue à celui de l'anglais *thin.*

La lénition est ancienne en celtique. On en relève des traces dans les inscriptions gauloises et c'est elle qui expliquerait le *v* de *Cévennes* en face du *m* de la désignation grecque, *kemenoi,* des mêmes montagnes. La lénition qu'on trouve dans les langues romanes de l'ouest présente des traits qui font penser qu'elle a pu être amorcée par les habitudes de gens de langue celtique lorsqu'ils ont appris le latin. Mais on retrouve la lénition un peu partout, en hébreu et en finnois, par exemple. Ce qui, en la matière, caractérise en propre le celtique, ce sont les traces qu'elle a laissées dans la grammaire des différentes langues, une fois le processus phonétique réalisé. Le français, et ce n'est sans doute pas un hasard, présente des traces de phénomènes analogues sous la forme des liaisons : le latin *(il)las feminas* « ces femmes » a abouti en français à *les femmes* qui, à date très ancienne, se prononçait conformément à la

graphie, c'est-à-dire [les femes] ; *(il)los infantes* aboutit à *les enfants*, prononcé autrefois [lezẽfãs], avec entre voyelles lénition du *s* en [z]. Aujourd'hui, le *s* est tombé dans *les femmes*, mais le [z] s'entend encore dans *les enfants*. Dans les langues celtiques, beaucoup de marques de flexion sont de ce type. Soit le mot irlandais *beatha* « vie », un des éléments du composé *uisge beatha* « eau de vie, whisky ». Il se prononce aujourd'hui [bʲɑhə]. La graphie, très archaïsante, reproduit la prononciation du XIIᵉ siècle. Placé devant ce mot, le possessif de 3ᵉ personne « son, sa, ses, leur, leurs » a la même forme, *a* [ə], mais selon que le possesseur est masculin, féminin ou pluriel, l'initiale de *beatha* va varier. S'il s'agit d'un masculin, « sa vie », angl. *his life*, sera *a bheatha* [ə vʲɑhə]. S'il s'agit d'un féminin, on aura *a beatha*, angl. *her life*, avec préservation de l'initiale, donc [ə biɑhə]. S'il y a plusieurs possesseurs, « leur vie » sera *a mbeatha* [əmʲɑhə], angl. *their life*. Le possessif, ici, représente, en fait, un ancien personnel au génitif, donc « de lui, d'elle, d'eux ». Ce génitif était, au masculin, marqué par *-ī*, comme c'est fréquemment le cas en latin. Le *b-* initial de *beatha* était donc entre voyelles et il s'est régulièrement affaibli en [β] (le *v* espagnol dans *la vaca*), puis est devenu un [v] ordinaire. Au féminin, le génitif était marqué par *-s* (comme dans le latin *pater familias* « père de famille »). Protégé par cet *-s*, le *b-* initial de *beatha* s'est conservé. Au pluriel, le génitif était en *-m*, comme dans le latin *hominum* « des hommes ». La succession *....m b....* a été traitée comme tous les *-mb-* de la langue, c'est-à-dire s'est simplifiée en *m-*.

De façon analogue, le mot breton pour « père », qui dérive d'un ancien **tata*, d'origine enfantine, d'où, vraisemblablement, l'anglais *dad, daddy* « papa », a les formes *tad, zad* ou *dad* selon le contexte grammatical.

Ce qui reste du celtique.

Entre les Romains au sud et les Germains au nord et à l'est, les Celtes ont vu leur domaine se réduire progressivement. On a longtemps pensé que la Gaule avait été intégralement romanisée avant l'arrivée des envahisseurs germains. On tend à croire aujourd'hui que le celtique s'était maintenu, sur le continent, dans des zones périphériques comme l'Armorique ou d'accès difficile comme la Suisse. Lorsque, sous la pression des Anglo-Saxons, les Celtes de Grande-Bretagne ont reflué vers l'ouest et que certains d'entre eux ont traversé la Manche, ils ont trouvé, dans ce qui allait désormais s'appeler la Bretagne, des populations qui parlaient encore le celte de Gaule. Les nouveaux arrivants, plus nombreux au nord qu'au sud, auraient imposé leur dialecte insulaire, sauf dans la région de Vannes où se seraient maintenues les formes continentales. Les trois dialectes du

breton dits KLT, c'est-à-dire Quimper (Kemper), Léon et Tréguier, accentuent, comme le gallois, l'avant-dernière syllabe. Le vannetais a l'accent sur la dernière.

Aujourd'hui, les parties de l'Occident européen restées celtophones sont généralement bilingues : franco-bretonne, anglo-galloise, anglo-écossaise. Dans l'Eire, où il a officiellement le statut de langue nationale, l'irlandais n'a pas réussi à s'imposer. Depuis le milieu du XIXᵉ siècle, où de terribles famines avaient décimé la population et où celle-ci avait émigré massivement en Amérique, l'usage de l'irlandais est resté limité aux populations rurales des franges occidentales de l'île.

LE TOKHARIEN

Il s'agit de la première des surprises que le XXᵉ siècle a réservées aux indo-européanistes. La découverte, au Turkestan chinois, et le déchiffrement de documents rédigés dans deux dialectes distincts, désignés le plus souvent au moyen des lettres A et B, n'a pas révolutionné la reconstruction indo-européenne : il s'agit d'une langue dont l'identification comme indo-européenne n'a guère fait difficulté, ce qui implique qu'elle n'apporte guère d'éléments nouveaux susceptibles de modifier notre conception de la langue commune. Elle est attestée à date assez récente, VIIᵉ siècle de notre ère, donc avec un vocabulaire assez renouvelé. Elle a surtout frappé les imaginations parce que, comme nous l'avons vu, on ne s'attendait pas à trouver une langue *centum* si loin à l'est. Sur les raisons qui ont pu amener ces gens jusqu'en cet endroit, on ne risque, semble-t-il, même pas des hypothèses. Les graphies employées, d'origine indienne, ont facilité le déchiffrement, mais elles sont médiocrement adaptées au système phonologique de la langue. Elles révèlent une évolution très particulière qui a abouti à confondre les trois séries de consonnes, disons les séries *t*, *d* et *dh*, en une seule, *t*. Ces confusions ont été en quelque sorte compensées par un phénomène analogue à celui que nous avons relevé pour le slave et qu'on retrouve, à l'autre bout du domaine, en irlandais. Il s'agit du transfert sur la consonne des caractéristiques de la voyelle suivante en même temps que la voyelle s'affaiblit. Lorsqu'elle disparaît, l'effectif des différentes consonnes se trouve multiplié, en principe, par trois. Soit trois formes *tati*, *tata* et *tatu*. L'articulation de *i* est anticipée dans le premier, d'où tat^{i}, celle de *u* dans le troisième, d'où tat^{u}, *tata* donnant *tat*. On obtient ainsi trois consonnes distinctes /t^{i}/, /t^{u}/, /t/ qui viennent, en quelque sorte, relayer les anciennes oppositions de la sourde **t* à la sonore **d* et à l'aspirée **dh*. Cette évolution que l'on reconstruit pour le tokharien est déjà

dépassée dans l'état de langue attesté par les documents, mais elle explique les curieuses mutations des voyelles (par ex. *keme* « dent » correspondant à grec *gomphós*) et leur répartition inattendue : noter la voyelle entre *p* et *n* dans *śpane* « sommeil » en face de grec *húpnos*, sanskrit *svapnas*.

LE HITTITE

La connaissance du hittite, déchiffré en 1917, a eu des répercussions beaucoup plus profondes sur la pratique comparatiste. C'était la langue d'un empire situé au centre de l'Asie mineure au deuxième millénaire avant notre ère. Les Hittites utilisaient la graphie akkadienne (ou, si l'on veut, assyro-babylonienne), d'origine sumérienne, qu'on qualifie de cunéiforme. Chaque caractère y était reproduit par un certain nombre de coups de stylet enfoncés dans des plaques d'argile. Il s'agissait d'un système mixte, combinant des idéogrammes et des syllabes. Les idéogrammes correspondaient, d'une langue à l'autre, aux mêmes notions : 8 est, si l'on veut, un idéogramme, compris également des Russes, des Allemands et des Français, même si pour les uns il s'agit de *vosem'*, pour les autres de *acht*, pour les troisièmes de *huit*. Ceux qui savaient lire le sumérien et l'assyro-babylonien n'avaient donc aucune difficulté à identifier les idéogrammes dans les textes hittites, sans pouvoir prononcer les formes correspondantes. En revanche, ils savaient prononcer sans les comprendre les éléments reproduits au moyen des signes syllabiques. L'hypothèse, qui s'est révélée correcte, que le hittite était une langue indo-européenne a permis, avec l'aide des idéogrammes du contexte dont on connaissait la valeur, d'identifier ces éléments prononçables.

Tout comme le tokharien, le hittite est une langue *centum*. Mais l'empire hittite étant situé beaucoup plus à l'ouest, ceci n'a pas paru trop étonnant. Ce qui a retardé l'identification immédiate de la langue à l'indo-européen tel qu'on l'avait reconstruit jusqu'alors a été ce qui apparaissait comme de profondes divergences, dues, en partie seulement, au fait que la nature de la graphie ne permettait pas d'identifier la forme de certains mots. En fait, tout se présente comme si la langue s'était séparée du tronc commun bien avant la séparation des autres branches de la famille : l'inexistence d'une distinction entre masculin et féminin, une économie assez particulière de la conjugaison où l'on identifie les mêmes éléments formels qu'ailleurs, mais dans des rapports différents, la préservation de certains des phonèmes que des considérations de type algébrique avaient amené à poser antérieurement dans la reconstruction en l'absence de toute attestation réelle,

tout cela met le hittite bien à part. Il n'est d'ailleurs pas le seul de son espèce. Il y a en Anatolie, vers la même époque, d'autres langues indo-européennes du même type, la mieux connue étant le louvite, attesté, comme le hittite, en graphie cunéiforme. D'autres, comme le lycien et le lydien, notées alphabétiquement, étaient celles de peuples bien identifiés dès l'antiquité. Leur interprétation pose des problèmes.

Il semble que le hittite ait distingué entre deux types d'occlusives, peut-être des sourdes et des sonores. Mais la graphie est telle que la distinction n'est marquée qu'à l'intérieur des mots où il y a une certaine constance dans les graphies doubles et les simples, les secondes correspondant généralement à des sonores des autres langues. En cette matière, comme ailleurs, les conditions dans lesquelles le hittite nous est connu ne permettent pas de se prononcer avec la même sécurité que dans le cas de langues de graphie plus homogène et moins ambiguë. Toutefois, le contact avec cet idiome très particulier a, jusqu'à un certain point, enhardi les chercheurs dans l'interprétation qu'ils faisaient des données d'autres langues.

BIBLIOGRAPHIE

Il serait vain d'essayer d'étayer chacune des affirmations ou des suggestions qu'on trouve dans ce chapitre par des références précises à un texte. On a affaire ici à une synthèse d'éléments fort disparates recueillis, au cours de soixante et quelques années, dans les textes ou les propos de mes maîtres et de mes collègues, mais repensés et réinterprétés dans mes propres termes. Sur un point particulier, comme, par exemple, celui des rapports mutuels des différents parlers germaniques, les opinions ont beaucoup évolué au cours de ce siècle, et l'on n'en finirait pas de citer tous les écrits et les enseignements qui ont pu influencer mon jugement, depuis les cours de mon maître Paul Verrier, à la Sorbonne, dans les années vingt, jusqu'à mes conversations avec mon collègue Witold Manczak à Chicago, il y a deux ans. Je ne citerai ici que les ouvrages que j'ai eu constamment sous la main pour vérifier formes et assertions, sans revenir sur les dictionnaires courants et sur ceux qui figurent dans les notes du chapitre 4, ci-dessus :

T. Burrow, *The Sanskrit Language*, Londres, s. d.

A. Leskien, *Handbuch der altbulgarischen Sprache*, 6ᵉ éd., Heidelberg, 1932.

A. Leskien, *Litauisches Lesebuch mit Grammatik und Wörterbuch*, Heidelberg, 1919.

Antoine Meillet et Jules Vendryes, *Traité de grammaire comparée des langues classiques*, 2ᵉ éd., Paris, 1948.

Holger PEDERSEN, *Tocharisch*, 2e éd., Copenhague, 1949.
Edgar H. STURTEVANT, *A Comparative Grammar of the Hittite Language*, Philadelphie, 1933.
Rudolf THURNEYSEN, *Handbuch des Altirischen*, Heidelberg, 1909.

CHAPITRE VI

DIVERGENCES, CONVERGENCES ET APPARENTEMENTS

Toute langue est en perpétuel état d'évolution. A cela il y a plusieurs raisons. La plus importante est que la structure des sociétés varie au cours du temps. Ce qui implique qu'il y a de nouvelles choses à dire ou qu'il faut présenter autrement ce qui existait antérieurement. Dans le monde contemporain où, du fait du raccourcissement des distances, les contacts entre les différentes cultures sont de plus en plus fréquents et intimes, il est à peine besoin de rappeler cette permanente instabilité des sociétés, même si elle n'est pas nécessairement vécue comme telle. Elle se ramène en effet le plus souvent à la satisfaction de nouveaux besoins. De ce fait, elle n'aboutit pas à des conflits et passe généralement inaperçue. Ce n'est guère qu'à l'occasion de réflexions rétrospectives qu'on prend conscience que le monde a changé, que tel fruit est de consommation courante aujourd'hui qui n'existait pas il y a vingt ans sur les marchés, que telle mode vestimentaire appartient désormais au passé, que telle vision du monde n'a plus cours. La plus grande longévité d'aujourd'hui facilite ces retours en arrière. Mais ils sont de toutes les époques, comme en témoigne la notion de *laudator temporis acti*, de louangeur du temps passé, qui n'est pas d'hier.

La prise de conscience de l'instabilité des cultures contemporaines n'entraîne pas automatiquement le sentiment que la langue en soit affectée. Comme les mots qui disparaissent graduellement de l'usage demeurent dans la littérature, les usagers de la langue n'ont pas le sentiment qu'il y ait jamais renouvellement du vocabulaire. Même s'ils prennent soudain conscience qu'ils vivaient, il n'y a pas si longtemps, sans connaître ni pratiquer le fruit de l'avocatier, le magret de canard, l'anorak ou les sicav, cela ne veut pas dire qu'ils perçoivent l'intrusion de ces nouveautés comme une atteinte à l'intégrité du lexique traditionnel. Même si l'on attire leur attention sur l'afflux de mots nouveaux, leur réaction, au demeurant assez saine, sera sans doute que

cette écume n'affecte pas la langue dans ses profondeurs, sa prononciation, sa grammaire, son lexique de base. Et cependant, si l'on s'écoute parler et qu'on essaye de le faire comme si l'on vivait encore en 1900, dans l'ignorance des deux guerres mondiales et de tous les changements, matériels, intellectuels et moraux qui se sont accumulés depuis lors, on constatera sans doute qu'un énoncé sur deux est incompréhensible.

Sans doute peut-on faire valoir que ceci vaut surtout pour notre siècle et qu'à côté des sociétés qui sont, pour ainsi dire, dans le mouvement, il y a celles, « primitives », qui restent, aujourd'hui encore, à l'âge de la pierre. Il est en effet vraisemblable que le rythme de l'évolution des langues, dans la mesure où il est un reflet de l'évolution du monde, est beaucoup plus lent dans des cellules sociales très homogènes et renfermées sur elles-mêmes que dans nos complexes nationaux et internationaux contemporains. Mais, même en dehors de tous contacts, une langue a quelques chances d'évoluer. Elle est en effet le lieu où sont perpétuellement en conflit les besoins de communication, les impératifs d'une tradition, la lenteur de certains esprits s'opposant à la vivacité des autres. Chez l'individu lui-même interviennent les asymétries des organes dits de la parole, qui, en fait, servent en priorité bien d'autres besoins.

Le jeu du maxillaire.

Rien n'illustre mieux les causes permanentes d'instabilité des langues au niveau le plus élémentaire, celui des sons du langage, que le jeu du maxillaire dans la production de ceux-ci. Les voyelles, ces éléments obligés de toute parole, ne sont pas autre chose que le son produit par les vibrations des cordes vocales qui se répercutent dans la cavité buccale. Ce qui les distingue les unes des autres est la forme assumée par cette cavité. Celle-ci dépend du jeu des organes qui s'y trouvent, de la poussée en avant ou en arrière de la masse de la langue, du jeu des lèvres, de la position du voile du palais, mais également, dans une large mesure, du degré d'abaissement du maxillaire inférieur qui conditionne l'ouverture de la bouche. Cette ouverture est considérable pour [a], faible pour [i] et [u] (*ou* français), moyenne pour [e] (*é* ou *è* du français) et [o]. Pour s'ouvrir, la mâchoire inférieure ne se meut pas exactement de haut en bas, mais, en fait, pivote sur ses apophyses, de telle sorte que pour un même angle d'ouverture de la bouche la distance d'un maxillaire à l'autre est plus grande à l'avant qu'à l'arrière. Or, l'observation des systèmes de voyelles des différentes langues montre qu'elles tendent à utiliser *les mêmes angles* des maxillaires, que la langue se masse vers l'avant ou vers l'arrière de la cavité.

Une langue qui distingue entre trois degrés d'ouverture vers l'avant

pour distinguer entre un /a/ (la voyelle de *patte*), un /e/ (celle de *clé*) et un /i/, a des chances de distinguer trois degrés à l'arrière, pour /ɑ/ (la voyelle de *pâte*), pour /o/ (celle de *clos*) et pour /u/ (celle de *clou*). Si maintenant une langue fait usage de quatre degrés d'ouverture vers l'avant et distingue entre /a/, /ɛ/ (français *è*), /é/ (français *è*) et /i/, on peut s'attendre à ce qu'elle réalise, à l'arrière, les mêmes angles des maxillaires et distingue entre /ɑ/, /ɔ/ (dans le français *sotte*), /o/ (dans le français *saute*) et /u/ (*ou* en français). C'est bien ce qu'on constate en français contemporain, en italien, en danois, et ailleurs. Là où, comme dans ces trois dernières langues, la distinction est moins constante entre deux voyelles à l'avant (par exemple *é* et *è* en français entre lesquels on hésite dans *antérieur, exact, quai, gai,* etc.), elle est également moins constante à l'arrière pour les voyelles correspondant au même angle d'ouverture ([ɔ] et [o] en français). Si une voyelle d'avant se diphtongue, c'est-à-dire que varie le degré d'ouverture au cours de l'émission, comme pour le [e^{i}] de l'anglais *pale*, par exemple, on constate, à l'arrière, pour le [o^{u}] de *pole*, la même tendance à se diphtonguer.

Toutefois, du fait de la configuration des organes, les distances entre les positions qu'ils assument pour les différents sons et, en conséquence, les différences entre les timbres réalisés seront moindres à l'arrière : dans beaucoup de langues, on trouvera une voyelle distincte de plus à l'avant : en allemand du sud, par exemple, où l'on distingue entre *segen* « bénir » et *sägen* « scier », on a, pour les longues, trois voyelles /ɛ:/, /e:/, /i:/ à l'avant, mais seulement deux voyelles /o:/, /u:/ à l'arrière, /a:/ restant neutre. Jusqu'au XVIIIe siècle, en français, on distinguait à l'avant entre *é* long (*idée*), *é* bref (*dé*), *ê* long (*être*) et *è* bref, (*mettre*) mais, à l'arrière, uniquement entre *ô* long (*saute*) et *o* bref (*sotte*). Cependant, l'examen de l'évolution des systèmes vocaliques les plus divers indique que la disproportion entre l'avant et l'arrière n'est pas de 4 à 3, mais intermédiaire entre 4/3 et 3/3, d'où une cause permanente d'instabilité : en termes un peu naïfs, quatre voyelles à l'avant c'est trop, trois voyelles ce n'est pas assez. L'histoire des voyelles du français au cours des trois derniers siècles est celle de l'extension des quatre degrés d'ouverture de l'avant à la série d'arrière, l'opposition de *ô* long à *o* bref passant à celle de *ô* fermé à *o* ouvert, d'où un rapport 4/4. En allemand, les usages du nord ramènent, pour toutes les voyelles, 4/3 à 3/3 en confondant dans la prononciation *segen* et *sägen*.

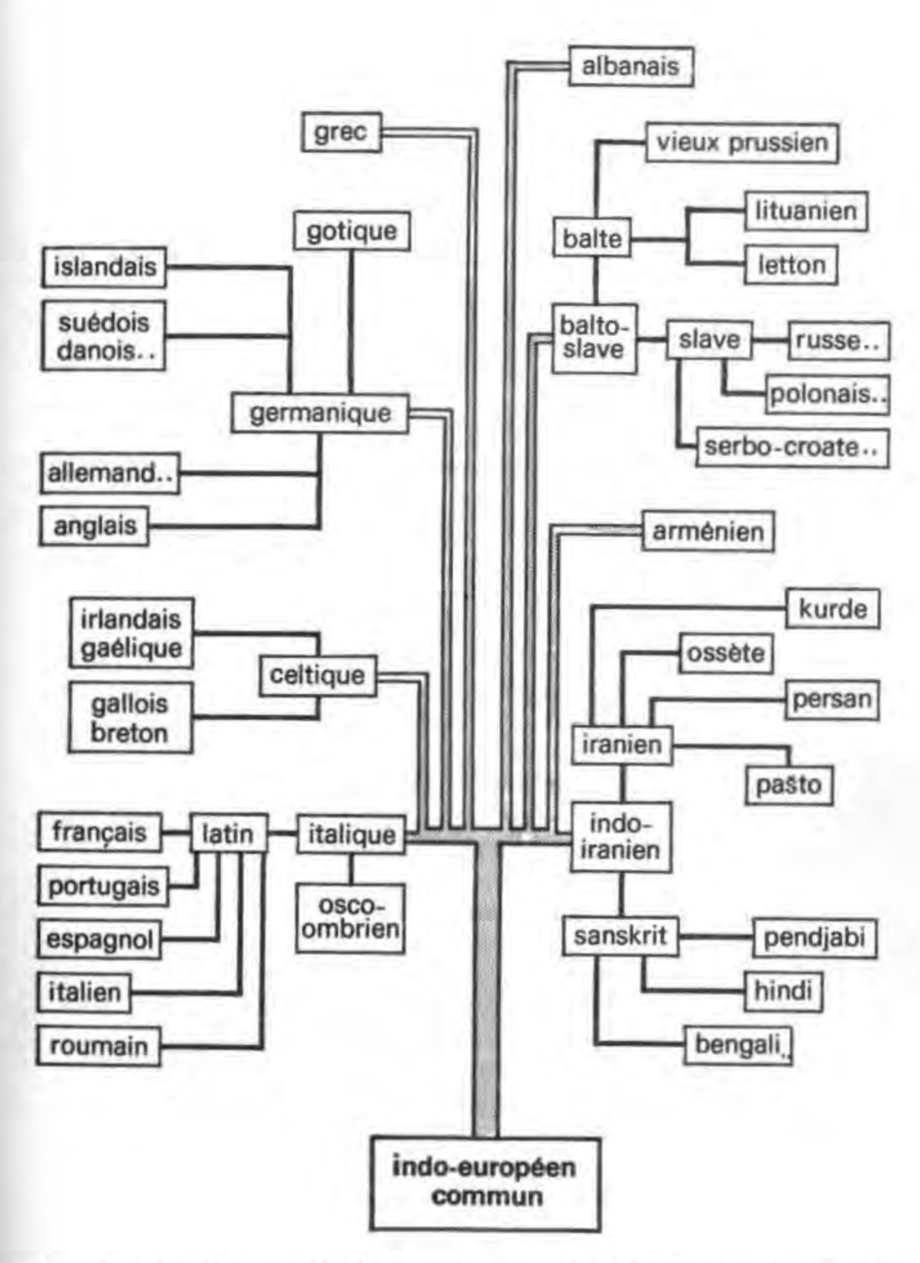
L'instabilité grammaticale.

Sur un tout autre plan, on retrouve des causes analogues d'instabilité. Quand, dans une langue, on distingue un passé d'un présent, on

sera légitimement tenté de distinguer également un futur et c'est bien ce qu'on constate lorsqu'on suit pendant assez longtemps l'évolution d'une langue. Mais cette logique un peu simpliste se heurte à la réalité psychologique qui est qu'on se représente bien et assez objectivement un passé qu'on a vécu, mais imparfaitement les événements à venir et qu'on tend à les envisager sous l'angle d'une obligation, d'une volonté ou d'un désir. Il est donc normal, et c'est ce qui se vérifie à l'examen, que la futurité pure soit moins fréquemment exprimée que le prétérit et que le système du futur, là où il existe, soit moins développé que celui du passé : en français, par exemple, où l'on distingue dans le passé, entre *il fut* et *il était*, on ne connaît, pour le futur, que le seul *sera*. Il y aura conflit constant entre, d'une part, la tendance à compléter les paradigmes, par exemple à partir de *il a fait*, à combiner avec le participe passé tous les temps et tous les modes du verbe *avoir*, ou, en anglais, avec la forme en *-ing*, tous les temps et tous les modes du verbe *to be*, et, d'autre part, la résistance contre une expansion qui contredit jusqu'à un certain point les besoins réels de la communication langagière.

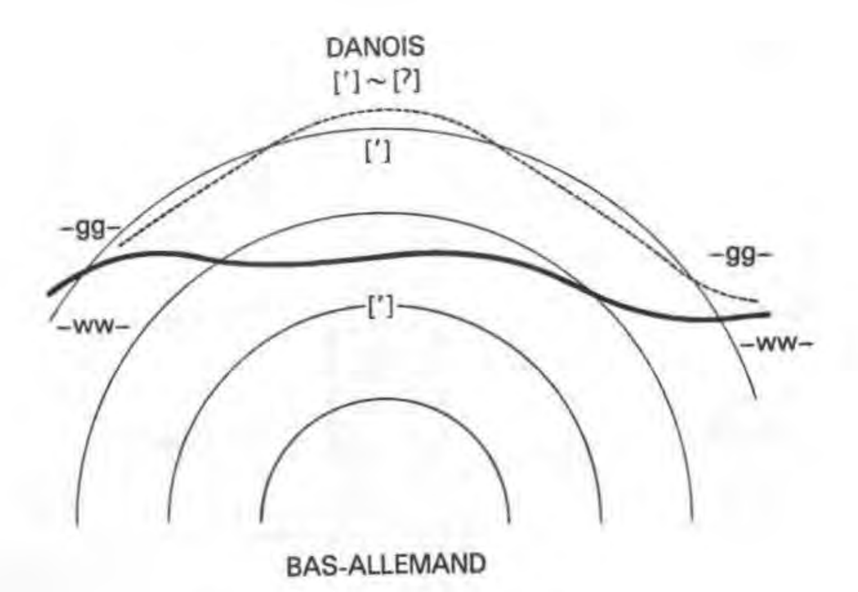
Le conditionnement interne.

Une implication importante de ce qui précède est que l'évolution d'une langue n'est pas uniquement déterminée par l'apparition de nouveaux besoins communicatifs qu'il faut satisfaire, mais que *la nature même du système linguistique*, à un moment donné de l'évolution, *va, dans une large mesure, en déterminer le sens :* une langue qui distingue un temps passé du temps présent a plus de chances de se créer un futur qu'une autre qui ne le fait pas ; une langue qui distingue un *p* aspiré (/p^h/) d'un *p* non aspiré a quelques chances d'étendre cette distinction à *t* et à *k* parce que ceux qui la parlent savent comment combiner une fermeture dans la bouche ([t] par exemple) avec l'articulation ouverte de la glotte, c'est-à-dire [h]. Ceci veut dire que lorsqu'un peuple se scinde en plusieurs branches qui, en émigrant, vont perdre contact, on peut s'attendre à ce que l'évolution dans toutes les branches soit au départ parallèle, un certain degré de parallélisme pouvant même continuer à s'affirmer lorsque les divers groupes seront soumis à des influences différentes. S'il s'agit de processus préhistoriques qu'on cherche à dater, on ne pourra faire coïncider la séparation avec le moment où l'on relève des divergences dans les évolutions respectives. On a, par exemple, essayé de reconstruire, par la comparaison des langues romanes, un roman primitif, celui qui fait abstraction de toutes les divergences que manifestent les langues issues du latin. Le résultat de cette opération ne se confond nullement avec le latin, car toutes ces langues ont commencé par évoluer dans le même



Cet arbre généalogique reflète les connaissances et les opinions généralement admises vers la fin du XIXe siècle. N'y figurent ni le tokharien, ni le hittite. Les points de suspension rappellent les langues qui n'ont pu être mentionnées faute de place, comme le néerlandais ou le tchèque.

L'arbre généalogique (Stammbaumtheorie)



Théorie des ondes (Wellentheorie)

Des déplacements de population avaient amené en contact deux « branches » différentes de Germains, ceux qui avaient changé *-ww-* en *-gg-* et ceux qui avaient conservé *-ww-*. Plus tard, l'influence du bas-allemand où n'existe qu'un seul type d'accent (noté [ˈ]) a affecté des parties du domaine danois en réduisant à un seul accent [ˈ] l'opposition danoise de deux accents, [ˈ] ~ [ˀ] (*mor* « mère » [moˈʁ] ~ *mord* « meurtre » [moʔʁ], le deuxième caractérisé par une petite toux).

sens. Si, par exemple, on fait intervenir le roumain, l'italien, le français, l'espagnol et le portugais et qu'on reconstruise par la comparaison le mot pour *cent,* ce n'est pas le *centum* ['kentum] latin qu'on retrouvera, mais quelque chose d'analogue à la forme italienne ['tšento].

La divergence et le Stammbaum.

Pendant longtemps, les linguistes qui étaient d'ordinaire, à ces époques, des comparatistes, n'ont envisagé l'évolution des langues que sous la forme de la divergence. Soit une langue parlée par un peuple. Ce peuple se scinde en plusieurs peuples distincts. La langue de chacun de ces derniers évolue selon son propre rythme et ses propres modes. Puis cette langue particulière se scinde à son tour en d'autres langues qui répètent la même opération. Tout cela peut être visualisé sous la forme d'un arbre généalogique, où le tronc représente la langue de départ d'où partent un certain nombre de branches correspondant à autant de langues. Celles-ci se divisent de nouveau en langues particulières et ainsi de suite. Pour l'indo-européen, le tronc va représenter la langue commune. On pourra d'abord envisager une fourche avec d'un côté les langues *centum*, de l'autre les langues *satem*. La branche *centum* se divise bientôt entre une branche hellénique, une branche italique, une branche celtique et une branche germanique. Cette dernière donne trois rameaux, celui des langues de l'est, celui des langues nordiques et celui des langues westiques. Ce dernier se scinde en une branche allemande et une branche anglo-frisonne. Cette dernière va donner un rameau frison et un rameau anglais. Celui-ci, de son côté, se ramifie en dialectes comme le kentien, les saxons de l'est, du sud et de l'ouest, le mercien et le northumbrien. Si l'on désirait poursuivre, il faudrait signaler qu'après quelques flottements, une seule de ces branches, le mercien, s'est imposée sous la forme de l'anglais du Moyen Age et de l'époque contemporaine. Il faut toutefois signaler qu'à l'époque où l'arbre généalogique — on disait à l'allemande le *Stammbaum* — était à la mode, on n'envisageait guère les dépérissements de cette sorte. On remarquera qu'il n'a pas été fait mention ci-dessus des langues tardivement identifiées comme le tokharien et le hittite, car la vision un peu simpliste du *Stammbaum* se trouvait dès lors dépassée.

La Wellentheorie.

La théorie qui, vers la fin du XIXe siècle, est venue, sinon réellement détrôner, du moins utilement supplémenter l'arbre généalogique est

celle des ondes (en allemand *Wellentheorie*). L'idée de base est que les changements linguistiques résultent d'innovations qui apparaissent en un point du domaine et s'étendent de proche en proche à la façon des vaguelettes concentriques qui se déplacent à partir du point de la surface de l'eau où est tombé un caillou. On voit immédiatement que l'optique a changé. Il ne s'agit plus de masses de migrants qui, quittant l'habitat primitif, s'enfoncent dans un désert humain. On envisage maintenant un domaine continu où, à partir de certains centres, vont se produire des dialectalisations qui, par additions successives d'innovations diverses, pourront aboutir à des idiomes parfaitement distincts. Si plusieurs innovations partent du même endroit et si les ondes qui en résultent s'arrêtent sensiblement aux mêmes points, il se créera là une frontière linguistique, c'est-à-dire que, du fait de la somme des innovations, les populations, de part et d'autre, auront des difficultés à communiquer linguistiquement. Si le processus se prolonge, on aboutira à l'incompréhension, donc à l'apparition de deux langues différentes.

Cette théorie des ondes apparaît au moment où l'on commence à s'intéresser activement aux dialectes, où l'on prend conscience que les grandes langues, celles qui ont un des noms qui figurent sur l'arbre généalogique, ne sont pas des données immédiates, mais résultent du nivellement linguistique d'un domaine hétérogène dû à une expansion politique, économique ou culturelle : le français est au départ le vernaculaire parlé à Paris qui, à travers les siècles, va s'imposer, par la force des armes ou le choix des marchands et des poètes, d'abord aux populations d'alentour qui modèlent leur usage sur celui de Paris, puis, au fur et à mesure qu'on s'éloigne de la capitale, à ceux qui l'apprennent comme langue commune et officielle tout en conservant, pour un temps, leur idiome local, parler d'oïl, parler d'oc, catalan, basque, breton, flamand ou alsacien.

La convergence.

Il est clair que, dans tous ses détails, ce modèle français ne saurait être retenu tel quel pour l'indo-européen du troisième ou deuxième millénaire avant notre ère. Mais il convient d'en retenir la notion de convergence, que celle-ci consiste à rapprocher son parler de celui de son interlocuteur de façon à faciliter la communication ou à se convaincre que, pour ce faire, la meilleure solution est d'apprendre sa langue. Il faut bien comprendre que la convergence, si elle n'affecte pas en même temps tous les sujets de la communauté, implique *ipso facto* divergence. Si elle amène un dialecte à se rapprocher d'un autre, elle a des chances d'impliquer une différenciation par rapport à d'autres voisins qui ne participent pas au rapprochement.

La recherche contemporaine a montré l'importance de la convergence sous sa forme la plus subtile, celle qui n'aboutit pas à l'abandon pur et simple d'une langue en faveur d'une autre. C'est pourquoi, en ce qui touche à l'indo-européen, on ne saurait écarter sans examen attentif de toutes les données, le point de vue généralement adopté par les chercheurs italiens au sujet des rapports entre le latin et les autres langues indo-européennes de l'Italie moyenne, selon lequel leurs ressemblances évidentes pourraient s'expliquer comme dues, non point à une communauté plus ancienne, mais à des contacts établis dans la péninsule. On a, ci-dessus, pris une position différente en faisant valoir que le consonantisme de toutes ces langues s'expliquait parfaitement à partir d'un état de langue où les phonèmes dits sonores aspirés avaient pris la forme de sourdes aspirées. Cela n'exclut nullement la possibilité d'une convergence. Toutefois cette convergence devrait être envisagée antérieurement à l'arrivée de tous ces peuples en Italie, dans une zone de l'Europe moyenne, disons au niveau du Danube, où les diverses variétés d'indo-européen qui y ont séjourné à une certaine époque ont, en commun, fait passer à des sourdes aspirées : *ph*, *th*, etc., des phonèmes qui apparaissent sous des formes sonores : *b*, *d*, etc., en celtique et en germanique. Si l'on part de la notation usuelle des phonèmes de cette série, à savoir, pour les labiales, par exemple, *b^h, on résumera ce qui précède en disant que dans les dialectes indo-européens parlés plus au nord, le celtique et le germanique, l'aspiration a disparu, d'où *b* ou [β], alors que dans cette zone de l'Europe moyenne l'aspiration est demeurée, mais l'occlusive s'est dévoisée, d'où **ph*. Ce **ph* est demeuré en grec classique pour ne devenir [f] que plus tard, alors que cette évolution s'est réalisée à date ancienne chez les populations qui avaient poussé jusqu'en Italie. On reviendra longuement ci-dessous sur les difficultés que pose l'identité des « sonores aspirées » en indo-européen commun.

Les Vénètes.

Le cas des Vénètes illustre bien la possibilité pour un même peuple du troisième ou du deuxième millénaire avant notre ère d'apparaître, plus tard, dans trois parties différentes du sous-continent, avec des formes de langues classées dans différentes branches de l'indo-européen.

On trouve quelques siècles avant notre ère, dans ce qui est aujourd'hui la Vénétie, autour de Venise (ital. *Venezia*), un peuple, les Vénètes (lat. *Veneti*), dont la langue est connue par quelques inscriptions. Un examen attentif permet de la rapprocher des autres langues indo-européennes de l'Italie centrale et notamment du latin.

D'autre part, au cours de sa conquête de la Gaule, César se heurte

en Armorique, dans ce qui est aujourd'hui le pays de Vannes, à un peuple de Vénètes (lat. *Veneti* également) dont personne ne met en doute l'appartenance à l'ensemble des cités gauloises. Dans cette région, désignée par les Romains comme la Vénétie, on parle aujourd'hui un dialecte breton, le vannetais. Certains estiment qu'à la différence des autres parlers d'Armorique qui proviennent de Grande-Bretagne, le vannetais pourrait être ce qui reste du gaulois sur le territoire de la France contemporaine. Il ne peut donc faire de doute que les Vénètes de la côte atlantique aient fait partie des Celtes.

A signaler, d'autre part, qu'une des branches du lac de Constance, au nord-est de la Suisse, portait dans l'antiquité le nom de *Venetus Lacus,* ce qui pourrait y indiquer la présence, à une certaine époque, de *Veneti,* sans qu'on puisse savoir si l'on doit les rattacher aux Vénètes de Gaule ou à ceux d'Italie.

Huit siècles plus tard, on trouve, dans les parlers germaniques, des formes *Winida* en haut-allemand, *Winedas* en vieil-anglais, qui correspondent régulièrement au *Veneti* [weneti:] latin. Ils y désignent les Slaves établis à cette époque dans la plaine de l'Europe du nord, de la Poméranie au Mecklembourg. Le terme a survécu, sous la forme de *Wendes,* pour désigner, chez les Allemands, les populations slavophones du Spreewald, au sud-est de Berlin, qui se nomment eux-mêmes Sorabes. En Autriche, l'adjectif *windisch* qualifie les localités habitées traditionnellement par des slavophones.

Que conclure de l'existence d'un même vocable pour désigner un peuple italique, un peuple celtique et un peuple slave ? Pourrait-il s'agir d'une simple homonymie ? Cela n'est guère vraisemblable. L'identité formelle est d'autant plus frappante qu'une forme indo-européenne avec *e* dans deux syllabes successives est un peu inattendue. On peut donc poser l'existence d'un peuple de Vénètes (['wenet-] avec une désinence de nominatif pluriel *-os,* celle attestée par le vieil-anglais *Winedas,* remplacée en celtique et en latin par *-ī*). On trouve dans ce nom une racine *wen-* bien connue, celle de *Vénus,* d'all. *Wunsch,* angl. *wish* « désir », danois *ven* et suédois *vän* « ami », en suédois également « charmant », donc, sans doute, « les aimables ». Si c'est bien ainsi que les Vénètes se désignaient eux-mêmes, on aime à croire que cette opinion favorable était partagée par leurs voisins. Toutefois, les peuplades indo-européennes semblent avoir été moins tentées de jouer de leur charme que de faire étalage de leur efficacité et de leur puissance. Aussi pencherait-on à attribuer ici, à la racine *wen-*, sa valeur de « gagner, conquérir ». On pourrait placer les Vénètes, vers la fin du troisième millénaire ou au début du deuxième, dans la plaine de l'Europe du nord, quelque part entre la Baltique et les monts de Bohême ou un peu plus à l'est ou à l'ouest. On peut leur supposer, à l'époque envisagée, une certaine autonomie linguistique n'empêchant

pas la compréhension avec les peuplades indo-européennes avoisinantes.

A un certain point du temps, ceux d'entre eux situés le plus à l'ouest de leur domaine subissent l'attraction de certains de leurs voisins qui vont devenir les Celtes. Comme chez ces derniers, leur **p* va commencer à s'affaiblir pour finalement passer à un *h*-, attesté dans le nom de la Forêt Hercynienne (< **perk*[w]*us* « chêne »). Ils suivront les Celtes dans leur mouvement vers l'ouest et se confondront finalement avec eux.

D'autres Vénètes vont s'agréger aux peuplades qui pousseront vers le sud pour séjourner quelque temps dans le bassin danubien. De ce séjour ils retiendront les sourdes aspirées des Italiques et des Grecs.

D'autres, enfin, restent en place, qui ont pu être partiellement germanisés, perdant ainsi leur identité du fait de pressions venues du nord. Il en est toutefois resté assez pour que, chez les Germains, le nom de *Wendes* reste la désignation normale de leurs voisins orientaux. On a vu que, de façon analogue, les Germains, pour désigner leurs voisins du sud-ouest, avaient retenu le nom d'une peuplade celte, les Volques avec lesquels ils se trouvaient en contact. Les voisins orientaux des Germains vont être, plus tard, vers le V[e] siècle de notre ère, recouverts par la marée slave. A cette date, après des millénaires de divergence entre les langues *centum* de l'ouest et le slave *satem*, on ne peut guère supposer une graduelle adaptation de leur langue à celle de leurs envahisseurs, mais bien son élimination. Il reste peut-être, dans certains parlers slaves de l'ouest, des traces de l'adoption de la langue par des allogènes : en face des voyelles nasales ou dénasalisées des autres langues slaves, le polonais présente des successions voyelle + consonne nasale : la « dent », qui était, en vieux-slave *ząbŭ* (c'est-à-dire, à peu près [ˈzõbə]) et qui a donné *zub* en russe ou en serbo-croate, *zôb* en slovène, *zəb* en bulgare, se prononce en polonais [zɔmp]. Cela rappelle ce qui s'est passé en France lorsque les Méridionaux ont dû apprendre le français et qu'ils ont reproduit *canton* comme *can'tong*.

Les Germains n'en ont pas moins continué à utiliser le terme de *Wendes* pour désigner des voisins dont la langue, avant ou après les migrations, leur était également incompréhensible. De façon analogue, si *Welsh* continue, en anglais, à désigner les celtophones du Pays de Galles, l'allemand *Welsch* s'applique aujourd'hui aux romanophones de Suisse. Rappelons ici que les Slaves désignent les Allemands comme « les muets », ceux avec qui on ne peut communiquer au moyen du langage.

Germanique, celtique et italique.

Le cas des Vénètes attire l'attention sur ce qui a dû être, à une certaine époque, un étroit apparentement entre Germains, Celtes et Italiques. Les comparatistes français ont été tentés, pour des raisons tenant à l'histoire de leur pays, de rapprocher l'italique et le celtique. On trouve effectivement des innovations communes aux deux branches : le génitif singulier, marqué par *(-o)s* ou *(e)s*, pouvait poser quelques problèmes d'homonymie dans le cas des radicaux se terminant en *-o-*, avec leur nominatif en *-os*. Les solutions divergent de branche à branche, mais l'italique et le celtique s'accordent pour remplacer ici *-os*, *-es* par *-ī*. On a été, par ailleurs, très impressionné par le parallélisme italo-celtique dans la répartition des dialectes, de part et d'autre, en deux types, l'un qui conserve initialement le $*k^w$, l'autre qui le remplace par *p* : pour « quatre », on a d'une part, avec *qu-* ou *c-*, le latin *quattuor* et l'irlandais *cethir*, d'autre part, avec *p-*, l'osque *petora* et le gallois *pedwar* ; de même, pour l'interrogatif, le lat. *quis* et l'irlandais *cía* s'opposent à l'osque *pis* et au gallois *pwy*. Toutefois, des rapprochements de même type ne manquent pas entre le celtique et le germanique, d'une part, l'italique et le germanique de l'autre. On a signalé ci-dessus les traitements exceptionnels que présentent en latin *lupus* et *bos* avec *-p-* et *b-* au lieu des *-qu-* et *u-* attendus. Mais on trouve des aberrances du même type en germanique : on a *f-* dans le gotique *fidwor* pour « quatre », *-f* dans *wulf* pour « loup » au lieu du h^w qui correspond régulièrement à $*k^w$ dans le relatif-interrogatif gotique *h^was* (lat. *quis*) ou *ah^{w}a* « fleuve » (lat. *aqua* « eau »).

Les rapports entre celtique et germanique méritent un traitement particulier. Au-delà de diverses particularités communes qui impliquent l'appartenance ancienne à un même ensemble, il y a une série d'emprunts faits par le germanique au celtique à une date relativement tardive où les deux langues avaient suffisamment divergé pour qu'on puisse aisément identifier le sens du passage d'une langue à l'autre : il y a d'abord des termes se référant à l'organisation de la société, comme le gotique *reiks* (prononcer [ri :ks]), avec les valeurs de « chef » et de « riche », et son dérivé *reiki* désignant le domaine du chef, plus tard l'état et l'empire (all. *Reich*). La forme initiale est celle qu'on trouve dans le latin *rēx* « roi », au pluriel *rēgēs*. Seul parmi les langues indo-européennes, le celtique a régulièrement changé le *ē* long en *ī* long. En face du latin *Rhēnum*, le français et l'allemand ont respectivement *le Rhin* et *der Rhein* qui portent témoignage de l'ancien *ī* du celtique. Les « Rois du monde », c'étaient les *Bituṝgēs*. Accentué à la latine sur le *-ī-*, la forme a abouti à *Berry*, nom de la province ; accentué à la

gauloise sur le *-u-*, elle a donné *Bourges*, qui en désigne la capitale. Inutile de rappeler ici *Vercingétorix* qui n'est pas autre chose que « le chef suprême des guerriers », donc, peut-être, un titre plutôt qu'un nom. Un autre emprunt intéressant est le mot qu'on trouve en gaulois sous la forme (latinisée) *ambactus* qui désignait le serviteur ; en vieux-haut-allemand, la forme est *ambaht*, réduit, en allemand moderne, à *Amt* qui désigne le service public, un fonctionnaire étant un *Beamter*. Loin de se dévaluer au cours des siècles, le mot a suivi la valorisation de son équivalent latin *minister* vite spécialisé au sens de serviteur de Dieu ou du prince. C'est d'*ambactus* que dérive *ambassade*, *ambassadeur*. C'est dire quelles connotations favorables pouvaient, dès le départ, s'attacher à ce terme.

Dans un tout autre domaine, celui des techniques, il y a quelque chance pour que l'anglais *leather* et l'allemand *Leder* « cuir » soient empruntés au celtique, vieil-irlandais *lethar*, gallois *lledr* dérivé de la racine de lat. *pellis* « peau » avec chute régulière du *p-*, soit, au départ, *pel-* — qu'on retrouve dans l'anglais *film* (*p-*>*f-*) — alternant régulièrement avec *ple-* et suivi du suffixe *-tro-* bien identifié.

Nous verrons, ci-dessous, p. 255, que les Germains ont emprunté aux Celtes le nom du fer à divers stades de l'évolution de la langue.

La possibilité d'identification de tels emprunts implique que Celtes et Germains ont rétabli des relations après une période où s'étaient relâchés les contacts qu'on est tenté de supposer entre eux à date ancienne. Il convient, bien entendu, de ne pas s'imaginer de façon trop simpliste les rapports qui pouvaient exister au troisième millénaire entre ceux qui allaient plus tard représenter les différentes branches de la famille. Sur la foi de leur répartition ultérieure, on tendrait à placer les Italiques au sud, les Celtes au milieu et les Germains au nord, avec d'autres peuples, les Vénètes peut-être, flanquant tous ceux-ci à l'est. Mais qu'on se rappelle qu'il faudra encore deux mille ans pour que les Germains se décident à amorcer leur mutation consonantique, celle qui permet vraiment de les distinguer de leurs congénaires. Pour ne prendre qu'un seul exemple, il y a toute chance pour que le mot qui deviendra plus tard *father* en anglais, *Vater* en allemand, *athir* en irlandais, *père* en français, *padre* en italien se soit encore, à l'époque considérée, prononcé uniformément *patēr* a peu près comme on le faisait à Athènes au temps de Périclès.

CHAPITRE VII

COMPARAISON ET RECONSTRUCTION

Il pourrait être intéressant, sinon de reconstruire la préhistoire de l'entreprise comparative, au moins de dégager les préjugés qui en ont retardé l'apparition. Cette entreprise commence au moment où les Occidentaux découvrent l'existence du sanskrit et se convainquent que ses évidentes ressemblances avec leurs propres langues classiques, grec et latin, ne peuvent être le fruit du hasard. Mais il est clair que, bien avant d'établir des contacts intimes avec la culture indienne, certains esprits curieux, parmi les Européens, n'avaient pu manquer de relever les ressemblances entre les langues dont ils avaient connaissance. Sans aller très loin, les termes de parenté les plus fondamentaux comme *père, mère, frère* en français, *father, mother* et *brother* en anglais, présentent d'évidents parallélismes qui devaient frapper quiconque était prêt à se poser des questions en la matière. Mais, pour ce faire, il fallait au préalable se convaincre que les mots ne s'identifient pas avec les choses ou les êtres qu'ils désignent. La réaction première, en face des parallélismes formels du type de ceux que nous venons de citer, n'est pas qu'ils nécessitent une explication, mais plutôt que les relations exprimées par ces termes étant les mêmes dans les différentes communautés, il est bien naturel qu'on les désigne par des termes analogues. On serait, au contraire, prêt à accepter que ces formes soient moins différentes, voire identiques, de part et d'autre. C'est à ces réactions que prétend répondre le mythe de Babel. Le succès qu'ont rencontré naguère, auprès du public contemporain, les doctrines selon lesquelles toutes les langues seraient faites, au-delà de déviations purement superficielles, sur un modèle identique, démontre la permanence de ce désir — on serait tenté de dire ce besoin — d'identifier la réalité perçue et la langue.

La signification du parallélisme de *père, mère, frère, father, mother, brother* ne peut apparaître que si l'on a, sur son chemin, rencontré des

langues, senties, au premier contact, comme un peu scandaleuses, où un tel parallélisme n'existe plus. Une fois fait ce premier pas sur la route du dépaysement, on va s'attendre à ce que les divergences croissent en proportion de la distance, dans l'espace et dans le temps. On ne s'étonnera donc pas que le chinois diffère profondément du latin, puisqu'il se trouve à l'autre bout de la terre et que deux millénaires les séparent. Mais qu'on retrouve notre trio, sous la forme de *pitā(r), mātā(r), bhrātā(r)*, dans des parages aussi lointains que l'Inde, ne peut manquer de réveiller l'attention. On se demandera alors pourquoi des langues aussi ressemblantes se retrouvent situées si loin les unes des autres. Qui s'est déplacé ? Les Européens, d'Asie vers l'occident, ou les Indiens, d'Europe vers l'est ? C'est toute la préhistoire du monde connu qui se voit affectée par la réponse à cette question.

En fait, ce ne sont pas les Français et les Anglais, les premiers à s'être avisés des ressemblances linguistiques entre sanskrit et langues européennes, qui vont y voir la matière d'une discipline nouvelle, mais des Danois et des Allemands, à qui la possibilité de rapprocher leurs propres langues de celles, prestigieuses, de l'antiquité classique, apporte une importante satisfaction d'amour-propre. Si le Danois Rask est le premier à esquisser le cadre comparatif, ce seront, pourtant, pendant trois quarts de siècle, les Allemands qui, presque seuls, bâtiront l'édifice de la grammaire comparative indo-européenne. Le terme d'indo-germanique, qui reste usuel en Allemagne pour désigner cette famille de langues, est très indicatif du rôle qu'a joué la nouvelle discipline dans l'exaltation de la grandeur nationale. A une époque où cette discipline est à peu près inconnue ailleurs, toute université allemande qui se respecte offre une chaire de sanskrit où l'on s'affaire à rapprocher cette langue du latin et du grec, sans doute, mais aussi de l' « allemand » conçu comme recouvrant à peu près toutes les langues germaniques, du gotique à l'anglo-saxon et au vieil-islandais.

En France, c'est l'Alsacien Bréal et, un peu plus tard, le Genevois Saussure qui, dans le dernier tiers du XIXe siècle, acclimateront la nouvelle discipline. Il faudra attendre Meillet pour qu'au siècle suivant, Paris contribue de façon significative au progrès de la recherche comparative.

On peut légitimement se demander à quoi vise la comparaison des langues indo-européennes. S'agit-il d'une tentative pour pénétrer plus avant dans la préhistoire ? Cherche-t-on à construire le modèle qui, étendu à d'autres familles de langues, permettra de retrouver le passé linguistique de l'humanité ? Veut-on mettre à profit une documentation qui remonte beaucoup plus loin dans le temps que partout ailleurs, pour découvrir dans quelles conditions et selon quelles

modalités évoluent les langues ? Un peu tout cela, sans doute. Mais, en fait, la comparaison figure surtout, dans les programmes, comme annexe à des études philologiques visant à une meilleure compréhension des textes anciens. En anglais, *philology* a longtemps été la désignation officielle de la discipline. En France, elle reste généralement conçue comme la chasse gardée des philologues chevronnés. C'est là une façon de se prémunir contre les hypothèses hasardées et les divagations irresponsables, mais cela ne favorise pas la diffusion de ses acquis et n'incite guère à remettre en question les enseignements dépassés.

Tout commence, bien entendu, par le rapprochement de langues dont on suppose qu'elles dérivent d'une même langue plus ancienne. Le modèle des langues romanes, résultant toutes d'évolutions divergentes à partir du latin, n'a pas joué, en la matière, le rôle que l'on aurait pu en attendre. En fait, la comparaison résolue des langues romanes les unes avec les autres, loin de précéder celle des anciennes langues indo-européennes, l'a suivie dans le temps et n'a jamais suscité la vision dynamique de l'évolution linguistique qui aurait pu dériver de l'examen attentif de réalités langagières directement accessibles. Romanistes aussi bien qu'indo-européanistes, et plus peut-être qu'eux, sont restés, souvent jusqu'à ce jour, de purs philologues, fascinés par la permanence des textes, si sécurisante en face d'une parole insaisissable sans cesse ouverte à toutes les innovations que réclame la variété des expériences à communiquer.

Le rapprochement des langues se fait, bien entendu, point par point. Les ressemblances qui frappent tout d'abord sont entre les formes des mots, y compris celles qui se manifestent dans le radical, c'est-à-dire la partie du mot qui ne change pas, et celles qu'on trouve dans les finales, marques des cas, du nombre, des personnes, des temps et des modes. Au départ, on ne se montre pas trop regardant : on retient les identités formelles et l'on fait abstraction des déviations qui sont conçues comme caractérisant chaque langue particulière. Mais, très vite, on constate que ces déviations ne se font pas au hasard, qu'à un *p-* du sanskrit coïncide toujours un *f-* du gotique, non seulement dans *pitār* ~ *fadar*, mais dans *paçu* ~ *faihu* qui désigne le bétail, dans *pāt* ~ *fōtus* « pied », dans *pūrnas* ~ *fulls* « plein » et partout où la forme et le sens suggèrent un rapprochement. On aboutit donc à des correspondances régulières du type sanskrit *p-* ~ grec *p-* ~ latin *p-* ~ slave *p-* ~ got. *f-* ~ irlandais zéro.

On s'est bien entendu vite demandé quelles langues avaient innové et lesquelles avaient conservé la forme primitive. Dans bien des cas, la réponse ne pouvait guère faire de doute : le *p-* du sanskrit, par exemple, était nettement majoritaire ; les formes *f-* et zéro n'existaient chacune que dans une seule branche, le germanique d'une part, le

celtique de l'autre. A ceci s'ajoutait la vraisemblance phonétique : on ne pouvait concevoir un passage de zéro à [p] ; un durcissement de [f] ne pourrait aboutir à [p], mais un affaiblissement de [p] peut conduire à [f] par l'intermédiaire de [ɸ], frottement de l'air entre les deux lèvres qui, s'il doit rester nettement perceptible et, en conséquence, assurer sa fonction distinctive, doit passer à un [f] beaucoup plus audible. Si ce passage n'a pas lieu, c'est finalement le bruit de l'air passant dans la glotte, c'est-à-dire dans la partie la plus resserrée du chenal expiratoire, qu'on va percevoir, au lieu de celui du frottement labial. Dans la notation, [h] remplacera [ɸ]. Cet [h] lui-même a quelque chance de disparaître à son tour s'il peut le faire sans que se produisent des confusions.

Mais, en posant que *p* est la forme ancienne, on pousse au-delà de la simple correspondance. On reconstruit un phonème de la langue disparue, dont dérivent les langues attestées. Son caractère hypothétique se marque au moyen d'un astérisque préposé : donc **p*. Sur cette pente, rien n'empêchera qu'on reconstruise tous les phonèmes d'un mot donné : à partir de sanskrit *paçu*, gotique *faihu*, latin *pecu*, on reconstruira donc **peku*. Si maintenant l'on pense qu'on peut en faire autant pour tous les mots et toutes les formes grammaticales de la langue, cela voudra dire qu'on a reconstruit cette langue et que, par conséquent, rien ne saurait empêcher qu'on écrive, dans cette langue, un texte entier.

Il y a, à cela, un précédent fameux, celui de la fable composée par le comparatiste allemand August Schleicher et publiée en 1868 sous le titre d'*Avis akvasas ka*, c'est-à-dire « Le mouton et les chevaux ». On n'en citera que la première ligne : *Avis, jasmin varna na ā ast, dadarka akvams...*, « Un mouton auquel la laine n'était pas, vit des chevaux... ». Malheureusement, quelques années plus tard, certains principes fondamentaux de la reconstruction pratiquée jusqu'alors allaient être remis en question ; le mouton, d'*avis*, devenait **owis* et la laine, **wl̥nā* au lieu de *varna ;* aujourd'hui on poserait volontiers, respectivement, *H_3*ewis* et **wl̥*H_2*neH*$_2$. On s'explique sans doute que la tentative de Schleicher n'ait pas eu de suite et que beaucoup de comparatistes aient émis l'opinion qu'il valait mieux ne pas se risquer au-delà des correspondances. Dans le *Dictionnaire étymologique de la langue latine* de Meillet et d'Ernout, par exemple, on trouve très peu de formes à astérisque. On est tenté d'approuver cette réserve jusqu'au moment où l'on constate que ces formes représentent la façon la plus simple de noter les correspondances et qu'il faut avoir celles-ci constamment en tête pour pouvoir se passer d'un rappel.

Le véritable inconvénient qu'il y a à penser la grammaire comparée en termes de reconstruction, c'est qu'on en oublie que l'indo-européen a évolué au cours du temps. Ce qu'on pense reconstruire, c'est la

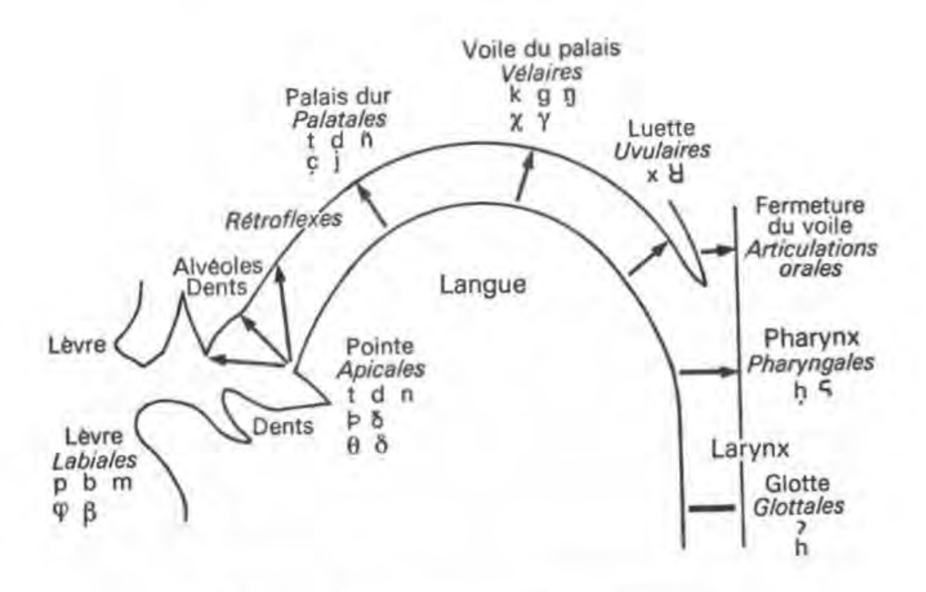


Schéma des articulations consonantiques

Les types articulatoires apparaissent en italique. Faute de place, ne figurent pas, sur le schéma, les variantes *apico-dentales* et *apico-alvéolaires* des *apicales*, non plus que les *labio-dentales* (f v). Pour chaque type, figurent, sur une première ligne, les occlusives, sur une seconde, les continues. Là où il n'y a qu'une seule ligne, il s'agit de continues. Pour les apicales, on distingue les continues d'articulation ferme (deuxième ligne) de celles d'articulation lâche (troisième ligne).

Pour les *sifflantes* (s z) et les *chuintantes* (š ž), l'air s'écoule par le sillon médian de la langue, le frottement étant réalisé entre un point de l'avant de la langue et un point de l'avant de la voûte.

Dans le texte de l'ouvrage, les *rétroflexes* sont signalées au moyen d'un point souscrit, et les *palatales* et les *palatalisées* indistinctement au moyen d'un accent aigu postposé. Dans les reconstructions, il s'agit de palatales. Dans la notation du russe, on a affaire à des palatalisées combinant une articulation spécifique et une articulation palatale.

langue parlée au moment même où le peuple de langue indo-européenne va se scinder en autant de branches distinctes que celles qui sont attestées plus tard, sans préjudice, bien entendu, de celles qui ont disparu sans jamais entrer dans l'histoire. Mais nous percevons bien aujourd'hui que ceci est une vue de l'esprit, parce qu'il n'y a aucune vraisemblance que toutes les langues indo-européennes résultent d'une diaspora unique et soudaine. Le déchiffrement du hittite a été, à cet égard, décisif. Que le hittite soit apparenté au grec et au sanskrit, la chose est évidente, mais il lui manque évidemment des traits, comme, par exemple, la distinction d'un masculin et d'un féminin, qui sont communs aux deux piliers traditionnels de la comparaison. Souvent les formes sont là, comme ce qu'on appelle le *-i* du « présent », mais elles apparaissent avec une répartition et dans des conditions assez différentes. On peut également se laisser influencer par l'allure inattendue des formes résultant de l'utilisation par les Hittites du syllabaire impliqué par la graphie cunéiforme. La réaction du hittitologue et comparatiste Edgar Sturtevant a été très indicative de la façon dont on comprenait, en fait, la reconstruction : l'indo-européen résultait de la projection sur un écran des produits de la comparaison des langues traditionnellement reconnues comme apparentées. Puisque l'adjonction du hittite à ces dernières aboutissait à brouiller l'image, il fallait envisager un autre écran où se projetterait l'image d'une autre langue reconstruite, l'indo-hittite.

Sturtevant espérait sauver ainsi l'image traditionnelle de l'indo-européen reconstruit. Mais cela en valait-il la peine à un moment où l'on était prêt à remettre en question certaines pratiques fondamentales ? En fait, il n'a guère été suivi et l'on ne parle plus de l'indo-hittite. Mais on aurait bien fait de mieux dégager et de mieux expliciter les raisons pour lesquelles on devait écarter les propositions de Sturtevant. Il aurait fallu, en fait, faire prendre conscience à tous les chercheurs concernés qu'il n'était plus question désormais de reconstruire un état de langue censé représenter la pratique de tout un peuple à la veille de son éparpillement à travers le monde. Ce qu'il fallait tenter d'imaginer était la dynamique d'une évolution linguistique, en remontant aussi loin que possible dans le passé et en extrapolant à partir des processus évolutifs identifiés par l'examen des états de langue successifs. Ces états de langue ne se limitaient plus à ceux pour lesquels on possédait des textes, mais comprenaient ceux qu'on restituait par la comparaison et ceux que suggérait une interprétation des archaïsmes attestés, soutenue par l'expérience fournie, depuis cinquante ans, par la description des structures linguistiques les plus diverses.

Dans ce qui suit, il ne saurait être question de pousser les recherches dans cette voie, ni même, simplement, de préciser le programme de

telles recherches. On cherchera simplement à montrer, parfois dans le détail, quelles ont été jusqu'ici les conceptions successives de la structure de l'indo-européen, celles-mêmes qui ont été sommairement illustrées par les trois versions successives de la reconstruction du mot désignant le mouton : **avis*, **owis*, *H_3*ewis*.

Bien qu'une structure linguistique soit un tout où chaque partie conditionne toutes les autres, on ne peut en comprendre le fonctionnement sans y distinguer différents plans qu'il convient d'examiner successivement, en commençant par ceux où l'on rencontre les unités qu'il est indispensable d'identifier avant d'accéder aux énoncés, c'est-à-dire aux formes langagières où s'organise l'expérience qu'on désire communiquer. On commencera donc par la phonologie, c'est-à-dire l'examen des phonèmes et plus généralement des traits qui permettent d'identifier la forme des éléments signifiants. Suivra l'examen des formes grammaticales, des considérations relatives aux structures fondamentales de l'énoncé et, finalement, un survol du vocabulaire.

CHAPITRE VIII

CHANGEMENTS PHONÉTIQUES ET ANALOGIE

La régularité des changements phonétiques.

La comparaison des langues indo-européennes et les efforts de reconstruction de la langue commune ont pris un caractère proprement scientifique à partir du moment où l'on s'est convaincu que les changements phonétiques se faisaient de façon régulière, c'est-à-dire que si, dans une langue particulière et à une époque déterminée, dans un mot donné, *a* se changeait en *e*, on devait s'attendre à ce que, dans tout autre mot, se produise le même changement. Les seules déviations admissibles étaient celles où le contexte était différent : dans le passage du latin au français, les *a* ne se sont changés en *e* que dans ce qu'on appelle les syllabes ouvertes, celles où, dans la syllabe, aucune consonne ne suivait la voyelle. Ailleurs, *a* est conservé : *mare* (*ma-re*) donne *mer*, mais *parte*(*m*), prononcé *par-te*(*m*), donne *part*. En syllabe ouverte, mais devant *n* ou *m* de la syllabe suivante, *a* aboutit à *-ai- : pane*(*m*) donne *pain*, *fame*(*m*) donne *faim*. Dans tout ces cas, nous avons affaire à des changements réguliers.

Là où l'on rencontre des déviations, elles peuvent n'être qu'apparentes, lorsque, par exemple, les clercs, influencés par la forme latine *clarus*, ont changé, dans la graphie, le régulier *cler* en *clair*. Mais elles peuvent avoir été déterminées par diverses influences : dans *fiancée*, *fian-* se prononce régulièrement à Paris comme une seule syllabe, exactement comme dans *confiance*, où l'on retrouve la même racine, celle de *fier*, *confier ;* mais on entend également une prononciation *non régulière*, qui traite *fiancée* comme si le mot était écrit *fillancée* où la prononciation porte la marque d'un rapprochement de ce terme avec *fille*. On dira, dans ce cas, qu'un rapprochement analogique de *fiancée* et de *fille* a empêché l'évolution régulière de [i] + voyelle en [j] + voyelle. Il faut bien comprendre que ce rapprochement analogi-

que n'est qu'une hypothèse, le processus lui-même ayant échappé à l'observation. Il en va probablement de même dans un cas comme celui du remplacement de *cler* par *clair*, sauf si nous avions le témoignage explicite de celui qui a décidé d'écrire *clair* avec *-ai-*. C'est pourquoi certains chercheurs ne se résolvent qu'à contrecœur à avoir recours à de telles suppositions.

Déviations expressives.

En principe, le sens des mots où apparaît un son donné n'est pas susceptible d'influencer l'évolution de ce son. Il y a toutefois des cas où, parce que celui qui parle se sent affecté par la valeur du mot qu'il prononce, il pourra articuler le mot avec une énergie particulière. C'est ainsi qu'il n'est pas rare d'entendre le mot *affolant* prononcé avec un [f] très allongé. Ceci n'a rien à voir avec la graphie du mot avec deux *f* : on ne constate rien de tel avec les mots *affaire* ou *affranchir*. Une telle prononciation allongée peut être imitée et reproduite dans des circonstances où rien ne la justifie, c'est-à-dire où celui qui parle n'est pas particulièrement « affolé ». Ceci peut aboutir à affecter la prononciation du mot de façon permanente. Soit, par exemple, le mot anglais *knock* « frapper ». En vieil-anglais, il était orthographié *cnocian*, avec un *-c-* unique entre voyelles correspondant à un [k] de la prononciation. Si le mot avait évolué sous cette forme, il rimerait avec les mots *oak*, *soak*, *coke*. En fait, il devait, dès l'époque la plus ancienne, être prononcé avec un [k] allongé ou géminé et, de fait, quelques siècles plus tard, en moyen-anglais, la forme est attestée dans la graphie avec deux *-kk-* entre voyelles, donc *knokke*, ce qui explique sa prononciation d'aujourd'hui avec une voyelle brève.

Les linguistes de la fin du XIX[e] et du début du XX[e] siècles, qui avaient eu à lutter contre leurs devanciers pour maintenir le principe que l'évolution des sons de la langue n'était pas affectée par la valeur des mots dans lesquels ils apparaissaient, ont souvent nié la possibilité des allongements expressifs.

Pour expliquer des cas comme celui de *knock*, ils avaient inventé un suffixe en *-n* qui aurait été, antérieurement aux textes attestés, ajouté au radical de certains mots, par exemple à *knok-*, donnant un groupe *-kn-* qui serait plus tard réduit à *-kk-*. Mais comme on rencontrait en fait des groupes *kn*, notamment à l'initiale du mot ici en cause, il avait fallu inventer des conditions particulières : *-kn-* n'aurait passé à *-kk-* qu'entres voyelles et si l'accent portait sur la voyelle suivante. Mais comme, en germanique, l'accent ancien, de place variable, avait, à une date préhistorique, été remplacé par un nouvel accent placé automatiquement sur la première syllabe du mot, aucune vérification de la validité de cette « loi phonétique » n'était possible.

En fait, il est impossible de nier que la valeur du mot puisse, dans certains cas, affecter le sort des sons qui le composent. Mais cela veut dire simplement que, dans certains cas, cette valeur participe au conditionnement de l'évolution. Ce n'est donc pas n'importe où que pourront se fixer des allongements expressifs, mais là où les conditions d'emploi du terme en apportent une justification.

Pourquoi des « lois phonétiques » ?

Pendant longtemps, ceux-là même qui étaient les plus ardents défenseurs des « lois phonétiques sans exceptions » se demandaient comment on pouvait expliquer leur existence. Il a fallu arriver jusqu'aux années trente de ce siècle pour trouver la réponse à cette question : chaque langue a, en fait, un nombre déterminé d'unités phoniques qu'on désigne comme des phonèmes. A chaque phonème correspond une habitude articulatoire particulière. C'est pourquoi l'apprentissage d'une nouvelle langue nous pose, en la matière, des difficultés. S'il s'agit par exemple de l'anglais, le premier phonème, noté /Þ/, de *thin* réclame que l'air qui sort de la bouche passe par une fente réalisée entre la tranche antérieure de la langue et celle des dents d'en haut. Un Français n'a jamais appris à réaliser cette articulation. S'il place les organes en cause dans la position décrite, il ne sait que les mettre en contact, fermant ainsi le chenal expiratoire et finalement réalisant son propre phonème /t/ lorsque l'articulation se relâche. S'il veut réaliser le frottement qu'il perçoit lorsqu'il entend un [Þ], il poussera la langue plus en avant, en en tendant les muscles, et fera passer l'air dans un chenal produit par la dépression du sillon médian de l'organe, ce qui est sa façon de réaliser son propre phonème /s/. Le /Þ/ est donc une habitude qu'acquièrent les enfants qui apprennent l'anglais. Il n'est pas exclu que cette habitude se modifie insensiblement d'une génération à une autre, mais comme elle est une habitude indépendante du sens des mots dans la prononciation desquels elle se manifeste, il est normal que toute modification aille dans le même sens, quelle que soit la signification du discours. Là où il peut se produire des déviations, c'est lorsqu'une certaine habitude articulatoire entre en conflit avec celles qui la précèdent ou la suivent dans la parole.

Influence du contexte.

Soit un phonème /a/, en français, par exemple. C'est une habitude articulatoire qui combine les vibrations des cordes vocales, une bouche grande ouverte, pour réaliser une cavité maximale, et le relèvement du voile du palais. Soit un phonème suivant, /d/, par exemple, dans

adieu ; /d/ est une habitude articulatoire qui combine les vibrations des cordes vocales et le relèvement du voile du palais, ce qui n'implique jusqu'ici aucun conflit avec /a/. La différence est la fermeture de la bouche au niveau de la langue et des dents qui succèdera à la grande ouverture pour /a/. Cette succession brusque d'ouverture et de fermeture du chenal est la condition même du bon fonctionnement de la parole et il faudrait des circonstances très particulières, qui n'existent pas en français, pour que l'ouverture de /a/ vienne affecter la fermeture de /d/. Si la consonne qui suit /a/ est le phonème /n/, nous aurons affaire à une habitude articulatoire qui combine les mêmes mouvements que /d/, sauf que le voile du palais va, cette fois-ci, s'abaisser de façon que, l'air passant par le nez, le son s'accompagne de résonances nasales. Bien que normalement coïncidant avec la fermeture buccale, cet abaissement pourra être légèrement anticipé et c'est ce qu'on constate dans beaucoup de langues, l'anglais notamment, où devant /n/ ou /m/ la voyelle est partiellement, voire totalement, nasalisée : en Amérique, *Congress*, le Congrès, c'est-à-dire le parlement, s'entend comme ['kɑ̃ŋgres], voire ['kɑ̃gres]. Ce n'est pas le cas en français où, dans la succession /an/ de *la nuit* /lanɥi/ on prend, inconsciemment, bien garde de ne pas nasaliser le /a/ pour éviter toute confusion avec *l'ennui* /lɑ̃nɥi/.

Cette action des phonèmes les uns sur les autres dans la chaîne parlée a été relevée depuis longtemps, bien identifiée par les comparatistes et classée par les linguistes contemporains parmi les causes de variation. Nous avons à plusieurs reprises signalé la disparition de **p* en celtique et opéré avec elle. Elle se vérifie, en effet, dans la plupart des positions, entre voyelles (*uper* > *uer* > *wer*) tout comme à l'initiale. Mais on constate que dans un groupe *-pn-* le *-p-* s'affaiblit, certes, mais en se sonorisant pour aboutir à un [u̯] non syllabique qui fait diphtongue avec la voyelle précédente. C'est ainsi qu'on peut dériver l'irlandais *cúan* « port » d'un plus ancien **kopno* par les intermédiaires [ko : n] < [koun] < [koβno]. De ce **kopno* procèdent régulièrement le danois *havn*, l'anglais *haven* et l'allemand *Hafen*. Cette forme reconstruite **kopno-* illustre bien la façon dont les comparatistes ont procédé pour marquer le détail des correspondances de langue à langue. Entre l'irlandais *cúan* et l'all. *Hafen*, il n'y a plus que le *-n* final en commun, mais l'un et l'autre se rattachent à **kopno-* par des évolutions régulières, c'est-à-dire qui ne souffrent en principe aucune exception : tous les **k-* sont passés en germanique à *h-* et tous les *o* à *a* pour d'ailleurs tomber à la finale ; les **p* y sont passés à *-f-*, dans des conditions accentuelles déterminées ; le groupe final *-fn* a développé une voyelle d'appui en westique, d'où *-fen ;* en anglais, *-f-* entre voyelle est passé à *-v- ;* en danois, devant consonne, il est également passé à *-v-* pour finalement devenir un [u̯] non syllabique.

Variation de forme d'un même mot.

Si, comme on vient de le constater, un phonème évolue différemment selon le contexte où il apparaît, on peut s'attendre à ce qu'une même unité de sens apparaisse sous une certaine forme dans certains contextes, une autre forme dans d'autres. Soit une forme latine *amīcus* « ami » ; au pluriel, il apparaissait comme *amīcī* avec un *c* prononcé [k] comme au singulier. Au cours de l'évolution du latin à l'italien, [k] s'est palatalisé devant *i* et *e*, ce qui a fait passer *amīcī*, de [aˈmi : ki :] à [aˈmitši]. Ce pluriel étant fréquent, il s'est maintenu tel quel. Pour le mot *tronco* « tronc, tronqué », on attendrait un pluriel **tronci* avec [tš] également. Mais comme le pluriel des noms et adjectifs en *-o* se fait normalement en remplaçant simplement *-o* par *-i*, on a, par analogie avec la masse des pluriels où la forme du radical ne change pas, refait un pluriel *tronchi*, avec [k], comme dans *tronco*. En dépit de la graphie qui change, c'est *tronchi* qui est le pluriel régulier.

Il est fréquent que la différence de contexte soit représentée par la présence ou l'absence de l'accent tonique sur une syllabe donnée. En français, un *o* bref latin a donné régulièrement *eu* dans les syllabes accentuées, et *ou* dans les syllabes atones. C'est pourquoi on dit *il meurt*, avec *eu*, mais *nous mourons*, *mourir*, avec *ou* là où l'accent portait sur *-ons* ou sur *-ir* et non sur le radical. Comme les formes *meurt* et *mourir* sont toutes deux de grande fréquence, l'enfant qui apprend le français les entend l'une et l'autre et finira par répartir la forme en *-eu-* et la forme en *-ou-* à la satisfaction des adultes. Mais cela pourra prendre du temps, parce que, à partir du moment où il aura compris que *meur-* et *mour-* font référence à la même réalité, il sera tenté de les confondre. S'il vient d'entendre *l'oiseau va mourir*, il reprendra *s'il mourt...* Si son interlocuteur a prononcé *tu meurs*, il pourra rétorquer *vous meurez*. Dans tous ces cas, on dit que l'enfant opère par analogie. Ce jeu de l'analogie apparaît plus ou moins tôt chez l'enfant. Il est signe d'un progrès de l'intelligence. Il indique que l'enfant n'en est pas à répéter des phrases toutes faites, de façon souvent imparfaite, mais qu'il sait identifier les unités significatives, même lorsque leur forme varie d'un contexte à un autre.

L'analogie.

L'analogie joue un grand rôle dans l'évolution des langues, parce que chaque nouvelle génération d'enfants va tenter d'éliminer des formes que l'évolution phonétique régulière laisse attendre, mais qui compliquent inutilement le fonctionnement de la langue. En vieux-français, on ne conjuguait pas, comme aujourd'hui,... *il prouve, nous*

prouvons..., mais *il preuve* (écrit *prueve*), *nous prouvons*, comme *il meurt, nous mourons*. Comme le verbe *prouver* n'avait pas la même fréquence que *mourir*, l'analogie l'a emporté aux dépens des formes phonétiquement régulières qui ne subsistent plus que dans le substantif *preuve*.

Certaines circonstances peuvent favoriser l'action de l'analogie. En vieux-français l'alternance des voyelles valait également pour le verbe *laver*. On disait donc, non pas *il se lave, nous nous lavons* mais *il se lève, nous nous lavons*. Mais il y a eu vite conflit avec la forme (analogique) *il lève* du verbe *lever* qui a favorisé l'extension du *-a-* à toutes les formes de *laver*.

Le verbe *rêver* du français doit, en dernière analyse, être le même mot que l'anglais *rave* « divaguer », mais l'analogie a joué dans deux sens différents dans les deux langues.

On aperçoit sans doute comment le fonctionnement de l'analogie peut compliquer la tâche du comparatiste : les « lois phonétiques » devraient lui permettre de prévoir à coup sûr l'évolution d'une forme, mais le jeu de l'analogie est largement imprévisible. Il fait, par exemple, intervenir des fréquences dont le calcul peut être délicat dans une langue contemporaine à laquelle on accède directement, mais qu'on doit renoncer à déterminer dans un état de langue hypothétique. C'est pourquoi les esprits avides de rigueur formelle répugnent à faire intervenir l'analogie dans leurs raisonnements. On peut comprendre leurs réticences, mais comme l'importance de ses interventions dans le processus évolutif n'est pas niable, il faut lui restituer la place qu'elle occupait dans les opérations reconstructives de nos devanciers.

L'exemple du mot langue.

Pour illustrer le rôle décisif des analogies de tous ordres, nous allons considérer le cas du mot *langue*. Il se retrouve dans presque toutes les branches de la famille pour désigner l'organe et, très fréquemment, l'instrument de communication. Le prototype que l'on reconstruit a la forme *dn̥g^{h}weH_2 dans laquelle, des voyelles qui ont pu exister entre les consonnes, seule figure la dernière qui, comme nous le verrons plus loin, pourra à son tour disparaître. Il y a toute chance pour qu'il s'agisse, au départ, d'un composé. On ne connaît en effet pas de monème unique avec cinq consonnes. Par évolution régulière, on aboutit en latin à *dingua*, à peine attesté à côté de la forme normale *lingua*. Il y a, en latin, d'autres cas d'hésitation entre *d* et *l*, par exemple dans *olere* « sentir » en face de *odor*, exceptionnellement *olor* « odeur ». Mais, dans le cas de *langue*, il ne fait pas de doute que le *l* est dû à l'analogie du mot *linguo* « lécher ». Ce type d'analogie a reçu le nom d'attraction paronymique. Cette attraction peut jouer même en

l'absence de similitude de sens, comme lorsque *braquet*, qui désignait un sabre court, est devenu *briquet* par analogie avec l'appareil à produire du feu. A plus forte raison lorsque, comme ici, l'analogie sémantique et l'analogie formelle sont évidentes. Le *-a* final du latin est la réduction normale d'un ancien *-ā* long dérivé régulièrement de *-eH₂* à la finale. Comme nous le verrons plus loin, H_2 a pu se prononcer comme un *ach* allemand ou comme la *jota* espagnole. Sa prononciation rétractée entraîne une prononciation [a] de la voyelle qui s'allonge par compensation quand H_2 s'amuït.

Les formes germaniques, du got. *tuggō* (prononcer [tuŋgo:]) à l'anglais *tongue* et à l'allemand *Zunge*, sont parfaitement régulières.

La forme lituanienne *liežuvis* a un *l-* initial dû, comme en latin, à l'analogie du verbe pour « lécher » *liežiu*; le *-ie-* de la première syllabe reproduit également la forme du verbe; le traitement *-ž-* du g^h est régulier en lituanien (langue *satem*); le segment final *-uvis* est identique à celui de l'équivalent vieux-prussien *insuwis*, c'est-à-dire [inzuvis]. Dans cette forme, comme dans celle du vieux-slave qui va suivre, le *d* initial a disparu. Il peut aussi s'agir d'une analogie, mais on ne sait pas laquelle. Les mots baltes et slaves remontent à une forme où est tombée la voyelle qu'on a posée ci-dessus entre *-w-* et $-H_2$. Ceci n'a pu se produire que là où c'était la désinence qui était accentuée : en face d'un nominatif $\star$-g^hweH_2-s, on avait un génitif $\star$-g^huH_2-*és*. L'analogie a joué en faveur de cette dernière. Dans ce cas, *-w-* placé entre deux consonnes passe à *-u-*, d'où, pour le second élément, la forme $\star$-g^huH_2; l'accusatif est marqué après consonne au moyen d'un *m* syllabique, noté *-m̥*, qui passe à *-im* en slave et en baltique; par analogie avec cette forme, le radical est interprété comme étant en *-i* final, d'où le nominatif en *-is* de *liežuvis* et *insuwis*. Dans $-g^huH_2i$-, H_2 tombe régulièrement entre voyelles, mais il se développe, entre *-u-* et *-i-*, un [w] de liaison qui passera ultérieurement à *-v-*.

En slave, ce n'est pas la forme de l'accusatif qui va s'imposer par analogie, mais celle du nominatif, un nouveau nominatif, bien entendu, dérivé par analogie des formes accentuées sur la désinence, par exemple, d'un génitif $\star$-g^huH_2-*és*. On aura donc au nominatif : $\star$-$g^huH_2 + s$; dans ce contexte, H_2 s'est durci en *k* comme l'a fait le *ch* de l'allemand dans le mot *sechs* [zeks] de cette langue, d'où $\star$-g^huks; à l'accusatif, la marque *-m* va être ici traitée comme non-syllabique, par extension analogique des cas où le mot suivant commence par une voyelle : on avait, en baltique, $\star$-g^huH_2*m̥*, forme attendue devant, par exemple, *t* initial du mot suivant, un groupe $-H_2mt$- étant imprononçable; en slave, on peut reconstruire $\star$-g^huH_2m *o-*, avec le groupe $-H_2m$- parfaitement prononçable entre deux voyelles. La succession $-uH_2m$- passe régulièrement à *-ūm* où *-m* est la désinence de l'accusatif et *-ū-* la fin du radical. On a donc en présence *-uk-s* au nominatif, *-ū-m* à

l'accusatif. L'unification se fera en étendant le *-k-* à l'accusatif et le *-ū-* au nominatif, d'où *-ūk-s* et *-ūk-m̥*. Par analogie, ce *-ūk-* va s'étendre aux autres cas. Mais le remplacement des radicaux en consonne finale par des formes en voyelle, plus stables et favorisées parce que moins exposées aux traitements divergents du type de ceux que nous venons d'examiner, va changer notre *-ūk-* en *-ūko-*, avec un nominatif *-ūko-s*, un accusatif *-ūko-m*, un génitif en *-ūko-ot* > *-ūkōt*, etc. Le produit final serait donc **n̥g^{h}ūko-*. Mais, dès cette époque, *g^h* est passé à *z* (traitement *satem*), *n̥-* est devenu *in-*. C'est donc à partir de **inzūko-* que la forme va évoluer régulièrement : *in-* passe à *ī-*, puis *ē-* qui est attesté, puis en russe à *ia-* [ja] ; le *-ū-* perd son arrondissement comme tous les *ū* de la langue et la voyelle finale tombe, ce qui nous donne la forme russe contemporaine *iazyk*.

Nous devons renoncer à étendre nos analyses aux formes iraniennes et indiennes d'une part, aux formes celtiques d'autre part qui présenteraient, à côté de traits absolument réguliers, les mêmes types de déviation que nous avons relevés jusqu'ici.

Il est intéressant de noter que le genre féminin est attesté pour le mot *langue* partout où la voyelle a été conservée entre les deux dernières consonnes du radical, c'est-à-dire là où a été préservée la combinaison *-eH$_2$*, caractéristique de ce genre, ou le *-ā-* qui en dérive. Le nominatif dans ce cas n'est pas marqué par *-s*. En revanche, dès que la voyelle tombe, que le radical se termine désormais en *-i* comme en balte, en *-k* comme en slave, en *-hu* comme dans certaines formes iraniennes, c'est le masculin, genre non marqué, qui s'installe. Il n'est pas du tout question que la langue en elle-même soit considérée comme féminine. L'opposition de genre féminin ~ masculin est ici simplement une affaire de forme.

Le tokharien a une forme à métathèse, c'est-à-dire avec inversion des consonnes, qui est, dans le dialecte B, *käntwa*, alors qu'à partir de **dn̥g^{h}weH$_2$*, on attendrait *tänkwa*. L'arménien *lezu* doit également son *l* à *lizanem* « je lèche ».

La seule branche de l'indo-européen où n'apparaît pas un avatar de **dn̥g^{h}weH$_2$* est le grec, qui a un vocable distinct : *glôtta, glôssa*.

Les produits divergents de l'analogie.

Le mot *langue* illustre bien comment les jeux divergents de l'analogie, en combinaison, bien entendu, avec des évolutions phoniques régulières, mais variables de langue à langue, peuvent aboutir à des résultats aussi différents que le français *langue* et le russe *iazyk*. Dans tous les cas, pour « langue », l'analogie joue bien son rôle simplificateur. Dans chaque idiome, la forme obtenue présente un seul et même radical : en français *langue* a toujours cette forme, avec

simplement, dans la graphie, l'adjonction d'un *-s* au pluriel, trait qu'il partage avec les autres substantifs ; en latin *lingua* se décline comme tous les substantifs en *-a;* en russe *iazyk* est la base immuable à laquelle s'ajoutent si nécessaire, les finales casuelles caractéristiques des masculins.

Toutefois, il peut se produire que, pour un mot très fréquent, l'analogie ait, dans une même langue, joué dans des sens variables, si bien que l'adjonction des désinences va aller de pair avec le choix de radicaux différents. En français, par exemple, le verbe *aller,* auxiliaire dans le futur prochain *je vais aller,* ne le cède guère, en matière de fréquence, qu'aux autres auxiliaires *être* et *avoir.* Les enfants apprennent ses formes variables, *va, allons, ira, aille,* avant même qu'ils aient pu les identifier comme appartenant au même verbe : quelques formes analogiques des types de *on allera* ou *il faut qu'il alle* n'ont aucune chance de se fixer, parce que l'enfant qui, en écho à *nous allons,* produit *on allera,* connaît la forme *on ira* qui lui viendrait naturellement aux lèvres dans un autre contexte.

Le cas du mot *pont* et les formes qu'il assume en sanskrit illustrent bien cette possibilité. Nous avons déjà rencontré *pont* et signalé dans quelles circonstances ce mot a pris le sens de construction permettant le passage au-dessus d'un cours ou d'une nappe d'eau, alors qu'il n'avait au départ que la valeur de chemin. C'est bien le sens qu'a le mot en sanskrit. Au cours de sa flexion, il n'y présente pas moins de quatre radicaux différents : 1. *panthā-,* par ex. au nominatif singulier *panthā-s,* 2. *panthān-,* par ex., au nominatif pluriel *panthān-as,* 3. *path-,* par ex. au génitif sing. *path-as,* 4. *pathi-,* par ex. au datif pluriel *pathi-bhyas.* Les éléments de base se retrouvent dans 1. et 3. A partir du génitif *pathas,* on peut reconstruire le radical **pn̥tH-;* l'absence de voyelle dans le radical va de pair avec l'accent sur la désinence de génitif, **-e/os ;* c'est de là que dérive le *th* qu'on retrouve dans toute la flexion. Pour le nominatif, on peut reconstruire un radical **pn̥teH-,* la désinence étant *-s* sans voyelle ; ceci donnerait **patā-s ;* mais il faut poser une forme parallèle avec l'accent au début, donc **péntH-s ;* c'est celle qui est à l'origine du nominatif latin *pons* (< **ponts*) et qui explique le *pan-* du sanskrit. La forme attestée, *panthās,* combine donc le *th-* régulier au génitif où l'accent était sur la désinence, le *-ā-* régulier lorsque l'accent tombait sur la voyelle qui existait au départ entre *-t-* et *-H-*, et le *pan-* d'accentuation initiale. Le radical 2, avec un suffixe en *-n-* ajouté au radical précédent, est dû à l'analogie de *rājā* « roi », pluriel *rājānas.* Le radical 4, avec son élément *-i-*, est celui qui est attesté en vieux-prussien sous la forme *pintis* (< **pn̥tHi-*) et en slave, mais là avec une voyelle pleine entre *p-* et *-n-*, d'où *pǫtĭ* (< **pontHi-*). Le grec a assuré aux formes la stabilité en leur ajoutant la voyelle *-e/o-* comme l'a fait le slave pour le mot *jazyk.* A

partir de _*pn̥tHo-_ on a _pátos_ « chemin » ; à partir de _*pentHo-_, _póntos_ qui désigne la mer, chemin le plus direct entre beaucoup de cités grecques. On connaît le _Pont-Euxin_ comme désignation de la Mer Noire, mot-à-mot « la mer hospitalière », mais sans doute par antiphrase, car les Grecs craignaient les tribus barbares qui peuplaient ses bords.

Les présentations nécessairement condensées qui précèdent peuvent sans doute donner l'impression que les linguistes jonglent avec les formes qu'ils ont couchées sur le papier sans trop essayer de s'imaginer comment les locuteurs, souvent les jeunes enfants, mais aussi les adultes, sont amenés, dans l'usage qu'ils font de la langue, à choisir telle ou telle variante. Et, de fait, lorsqu'il s'agit de langues non attestées ou connues seulement par des textes, il n'est pas possible de découvrir, d'imaginer toutes les analogies qui ont pu jouer. A côté de la situation privilégiée où nous sommes, pour expliquer le _l-_ de _langue_, avec, partout où s'est produit le changement de _d-_ en _l-_, un mot du sens de « lécher » à _l-_ initial, nous rencontrons constamment, dans nos reconstructions, des divergences qui suggèrent qu'une attraction a dû avoir lieu, sans que notre connaissance trop lacunaire du parler en cause nous permette d'aller au-delà d'une hypothèse hasardée.

Comme partout ailleurs dans la recherche, il faut ici toujours être prêt à remettre une hypothèse en cause dès qu'interviennent de nouveaux éléments d'information. Mais ce serait se condamner à la stérilité que de se refuser à envisager toutes les hypothèses que suggèrent les connaissances que nous acquérons chaque jour de l'examen des réalités linguistiques les plus diverses. Après plus de cinquante ans d'une linguistique surtout descriptive, nous pouvons revenir à la reconstruction avec quelque chance de mieux comprendre ce qui s'est passé au cours des millénaires qui nous ont précédés.

CHAPITRE IX

LE SYSTÈME PHONOLOGIQUE

La phonologie est essentiellement l'étude des **unités distinctives** du discours dans chaque langue particulière. Ces unités distinctives sont désignées comme les phonèmes de la langue. Dans une graphie alphabétique, à chaque phonème correspond, en principe, une lettre. C'est ce que l'on constate, en français, dans les mots *calcul, joli, carnaval,* mais l'orthographe française comporte des entorses fréquentes et systématiques à ce principe, comme on le voit dans *roues* avec cinq lettres pour deux phonèmes. Tant qu'on opérait avec des langues anciennes où le principe était respecté, on pouvait, sans grands dommages, identifier lettre et phonème. Mais l'intérêt porté aujourd'hui aux langues les plus diverses impose la distinction.

On distinguera également la phonologie de la phonétique, cette dernière discipline traitant des sons du langage en général, la phonologie, elle, se concentrant sur le système des unités distinctives de chaque langue. Rappelons ici que les barres obliques marquent une notation phonologique, c'est-à-dire présentant les unités distinctives caractéristiques d'une langue donnée, par exemple /roš/ pour le français *roche*, alors que les crochets carrés de la notation phonétique attirent l'attention sur la réalité physique symbolisée par les signes qu'ils entourent, par exemple [ʁɔ̄š] pour certaines prononciations parisiennes du même mot.

A partir du sanskrit.

Lorsqu'il s'est agi d'identifier et de se représenter les phonèmes de l'indo-européen commun, l'image du sanskrit s'est, dans un premier temps, imposée absolument. On a même pu, au départ, envisager que cette langue s'identifiait totalement au sanskrit. Les premiers comparatistes étaient en fait dans l'émerveillement en face de l'harmonie et la

symétrie du système consonantique du sanskrit. Cette symétrie n'avait, en fait, rien d'extraordinaire et l'on en rencontre d'aussi parfaites dans beaucoup de langues, mais il se trouvait que les grammairiens indiens l'avaient bien mise en valeur. Avec ce que nous désignons aujourd'hui comme cinq séries de cinq ordres, l'essentiel du système se présente sous la forme du carré suivant :

Séries \ Ordres	I	II	III	IV	V	
1	p	t	ṭ	c	k	Sourdes
2	p^h	t^h	ṭ^h	c^h	k^h	Sourdes aspirées
3	b	d	ḍ	j	g	Sonores
4	b^h	d^h	ḍ^h	j^h	g^h	Sonores aspirées
5	m	n	ṇ	ñ		Nasales

Les ordres sont, de gauche à droite, I labial, II dental, III rétroflexe, IV palatal, V vélaire.

On s'est abstenu, pour ne pas affecter la symétrie, d'adjoindre les sifflantes dentale (*s*), rétroflexe (*ṣ*) et palatale (*ç*) qui forment une sixième série non représentée dans les ordres I et V.

A côté des consonnes, on trouve quatre « semi-voyelles », c'est-à-dire des réalités phoniques qui peuvent être syllabiques et fonctionner comme des voyelles, ou non syllabiques et fonctionner comme des consonnes. C'est le cas de *i* et de *u* qui deviennent *y* et *w* respectivement lorsqu'ils ne sont pas sommets de syllabe. C'est ce qui se passe en français lorsque le *i* de *je nie* ne fait plus syllabe dans *nous nions* dans la prononciation parisienne. Outre *y* et *w*, on trouve là, en sanskrit, *l* et *r* qui, effectivement, apparaissent comme des noyaux syllabiques, par exemple dans *pr̥thu-* « large », en deux syllabes. On trouve des faits analogues dans d'autres langues, en tchèque, par exemple, dans *prst* « doigt », *plný* « plein ». En fait, la graphie indienne distingue parfaitement les usages vocaliques et les usages consonantiques. Traditionnellement, ce qui est donné ici comme *w* est translitéré comme *v* en accord avec la prononciation d'aujourd'hui qui d'ailleurs existait peut-être dès l'époque où la graphie indienne a été mise au point.

Il y a essentiellement une voyelle qui est *a*, bref ou long. La langue connaît des *ē* et des *ō*, mais qui alternent avec *i* et *u* et sont encore sentis comme les *ai* et les *au* qu'ils étaient au départ. En français, la graphie, dans *aine* et *aune*, témoigne encore de l'existence ancienne d'une diphtongue pour ce qui se prononce, en fait, *êne* et *ône*. Il y a des « semi-voyelles » longues en fonction syllabique, donc des *ī*, des *ū*, voire des *r* et des *l* syllabiques longs. Mais tous ces phonèmes apparaissent comme les avatars de diverses combinaisons plus anciennes, et ceci suggère qu'il n'y avait, à l'origine, qu'une seule voyelle qui, en s'allongeant, en se réduisant à zéro, en se combinant diversement avec ses contextes, a donné la variété des vocalismes qu'on rencontre dans les diverses langues indo-européennes. Si, aujourd'hui, nous ne croyons plus à la primauté du sanskrit, nous sommes, plus peut-être que les premiers comparatistes, tentés d'y retrouver, sur ce point particulier, un modèle de ce qu'a pu être la langue indo-européenne à date très ancienne.

Du beau système des consonnes du sanskrit, on a vite exclu, pour la langue commune reconstruite, l'ordre n° III des rétroflexes. La connaissance, assez ancienne, des plus vieux textes iraniens, dénonçait clairement ces rétroflexes (on disait des « cacuminales ») comme une innovation indienne. Il restait donc les quatre ordres, numérotés ici I, II, IV et V, respectivement labial, apical, palatal et vélaire. La correspondance des deux premiers avec leurs équivalents des autres langues ne pouvait guère faire de doute : nous avons déjà rencontré skt. *pitā(r)*, grec *patēr*, lat. *pater ;* pour « trois » on a skt. *treyes*, grec *treis*, lat. *trēs*. Mais pour ce qui touchait au palatal et au vélaire, on constatait que les palatales du sanskrit, disons, par exemple, *j* dans *jñā-* « connaître », correspondaient à des vélaires, donc *g*, dans le latin *gnōscō* et le grec *gignoskō*. Quant aux vélaires sanskrites, leurs équivalents normaux, en latin et en germanique, étaient des labiovélaires, les interrogatifs étant en *k-* en sanskrit, mais en *qu-* en latin, en *hw-* en gotique. Comme il se trouvait toutefois quelques cas où à *k*, *g* du sanskrit correspondaient également *k*, *g* en grec ou en latin, par exemple *kravis* « chair crue », grec *kréas* « id. », lat. *cruor* « sang », *yuga-* « joug », gr. *zugón*, lat. *iugum*, on a posé, pour la langue commune, trois ordres différents, un ordre palatal avec une sourde notée **k'*, un ordre vélaire avec *k* et un ordre labiovélaire avec **k^w*. On tend à croire, aujourd'hui, qu'on n'y distinguait pas entre les palatales et les vélaires et que les cas où dans les langues *satem* on trouve, par exemple, un *k* pour un *k* grec ou un *c* latin, s'expliquent par des conditionnements et des développements analogiques particuliers. Nous ne ferons donc pas état, dans la suite, d'une différence ancienne entre palatales et vélaires et n'opérerons qu'avec deux ordres caractérisés, par exemple, pour les sourdes, par **k* et **k^w*.

C'est la reconstruction des *séries* consonantiques qui pose les problèmes les plus délicats. Les premiers comparatistes n'hésitaient pas, en la matière, à faire totalement confiance au sanskrit et reconstruisaient les quatre séries 1, 2, 3 et 4, de sourdes, de sourdes aspirées, de sonores et de sonores aspirées. Comme l'indien était la seule branche à présenter quatre séries, l'iranien lui-même n'en présentant que deux, on aurait pu penser à une innovation. Mais les quatre séries semblaient constituer une structure très cohérente et l'on fermait les yeux sur le caractère un peu marginal de la série 2 des sourdes aspirées. Nous y reviendrons ci-dessous. On continue d'ailleurs à trouver étrange un système où l'aspiration, qui va normalement de pair avec la sourdité, pourrait n'exister qu'en combinaison avec la sonorité.

Voici donc le système consonantique avec lequel on a opéré pendant tout le XIXe siècle, où l'on conserve trois ordres de dorsales et auquel certains continuent à se référer :

$^{*}p$	$^{*}t$	$^{*}k'$	$^{*}k$	$^{*}k^{w}$
$^{*}p^{h}$	$^{*}t^{h}$	$^{*}k^{h\prime}$	$^{*}k^{h}$	$^{*}k^{hw}$
$^{*}b$	$^{*}d$	$^{*}g'$	$^{*}g$	$^{*}g^{w}$
$^{*}b^{h}$	$^{*}d^{h}$	$^{*}g^{h\prime}$	$^{*}g^{h}$	$^{*}g^{hw}$

En ce qui touche à la notation, il faut signaler que beaucoup d'auteurs préfèrent les graphies du type $g^{w}h$ nées de l'idée un peu naïve que le *h* correspond à un élément postposé à l'articulation buccale, alors qu'en fait, aussi bien l'aspiration que la labialisation et l'occlusion, ici dorsale, sont des phénomènes concomitants, la successivité des éléments graphiques résultant de l'impossibilité où nous sommes de les combiner de façon plus intime.

On reviendra longuement ci-dessous sur ce qu'il faut penser de ce beau système.

La reconstruction classique ne connaît qu'une seule sifflante notée *s, les autres sifflantes du sanskrit étant l'une une innovation rétroflexe, l'autre le reflet régulier du *k', le *c* du tableau sanskrit résultant d'une palatalisation indo-iranienne particulière, dont les produits se sont, sauf en ce qui concerne la sourde simple, confondus avec ceux du processus *satem*. Cela rappelle ce qu'on trouve dans le passage du latin au français : on a [ž], aussi bien dans *gens*, de *gentes*, que dans *jaune*, de *galbinu(s)*, alors qu'on a [s] dans *cent*, de *centum*, mais [š] dans *champ*, de *campu(s)*.

On ne pose que deux nasales **m* et **n*, celles qui s'y ajoutent en sanskrit n'étant, au moins au départ, que des variantes, sans autonomie réelle.

On reconstruit deux « liquides », **l* et **r*, un **y* et un **w*. On retient, de ce que nous enseigne le sanskrit, que ces autres phonèmes

et les deux nasales, tous groupés sous la rubrique de sonantes, ont pu, en l'absence de voyelle, fonctionner comme sommet de syllabe. Ils sont, dans ce cas, notés *ḷ, *ṛ, ı, u, m̥ et n̥. Les groupes *al, *ar, *am, *an sont censés se comporter comme des diphtongues, au même titre que *ai, *au. On verra plus loin ce que cela implique.

En matière de voyelles, à côté de *a* et de *ā*, on retient, outre les sonantes en fonction syllabique, au moins deux longues *ī et *ū.

Enfin, on pose une voyelle *ə, dite *schwa*, lorsqu'un *i* de l'indo-iranien correspond à un *a* de toutes les autres langues, par exemple dans *pitā(r)* « père » en face de grec *patēr*, latin *pater*, v. irl. *athir*, got. *fadar*.

Une voyelle unique ?

Il faut arriver jusqu'au dernier tiers du XIXe siècle pour que certains traits fondamentaux de ce système soient mis en doute. Lorsqu'ils le sont, c'est qu'on s'est montré un peu plus exigeant en matière de comparaison. On se rappelle que, non content de participer à l'évolution *satem* qui fait passer *k' à une sifflante, l'indo-iranien palatalise parfois le produit [k] de l'ancien *k^w. Par exemple, l'équivalent de la particule *-que* « et » du latin n'est pas *ka*, mais *ca*, c'est-à-dire [kya] ou [tša]. En étudiant le conditionnement de cette palatalisation, on arrive à la conclusion qu'elle ne s'est produite que dans les contextes où le témoignage d'autres langues, comme le grec et le latin, laisse attendre un *e*, mais non le *o* qui alterne avec lui, par exemple dans *tego* « je couvre » en face de *toga* « la couverture, la toge ». Nous reviendrons longuement, ci-dessous, sur cette alternance. Ceci amène à la conclusion que la distinction entre *e* et *o* est ancienne, qu'elle s'était maintenue aux premiers temps de l'unité indo-iranienne et que la confusion des deux voyelles en *a* a été le résultat d'une évolution ultérieure propre à l'indo-iranien.

Pour bien comprendre ce dont il s'agit ici, il faut préciser ce que peut impliquer, trois mille ans avant notre ère, deux millénaires et demi avant le siècle de Périclès, ce que nous avons désigné comme la distinction entre *e* et *o*. En grec attique, le *é* de *légō* était à peu près celui du français *léguer*, et les *o* de *lógos* celui de *rôle*, c'est-à-dire deux réalisations vocaliques foncièrement différentes. Mais il ne peut faire de doute qu'elles remontent l'une et l'autre à la même voyelle. L'expérience que nous avons des langues nous indique qu'il s'agissait alors d'une voyelle que nous noterions *a* si nous l'entendions, mais dont l'articulation pouvait varier selon le contexte, un peu comme, à Paris, le *a* de *tasse* varie selon les gens, les uns faisant rimer le mot avec *fasse*, d'autres avec *lasse*. A l'époque où les ancêtres linguistiques des Grecs se sont séparés de ceux qui allaient devenir les Indo-Iraniens, on

peut envisager dans la langue deux timbres vocaliques peu différents, l'un analogue au *a* anglais de *cat*, noté [æ], l'autre à celui du *o* de *pot*, qu'on peut noter [å] pour l'usage britannique et [ɑ] pour l'usage américain. Dans ce qui va devenir l'indo-iranien, un [k] placé devant [æ] va passer à [kʲæ] et, plus tard, à [tšæ]; devant [ɑ], ce [k] ne bougera pas. Un peu plus tard, l'indo-iranien confondra [æ] et [ɑ], un peu comme les jeunes Parisiens d'aujourd'hui sont en train de confondre *patte* et *pâte* qui, il y a cinquante ou soixante ans, s'articulaient, dans les faubourgs, respectivement comme [pæt] et [pɑ̊t]. En grec, au contraire, la différence va s'accentuer, le [æ] passant à [ɛ] puis à [e], le [ɑ] à [ɔ] puis à [o].

Il n'est toutefois pas indifférent que la voyelle d'arrière ait eu, dans un dialecte donné, un timbre [å] ou un timbre [ɑ]. Là où le timbre était [å] il y avait place, entre [æ] et [å], pour un timbre [ɑ] attesté dans une création familière comme **kaput* « tête » désignant, sans doute, au départ un pot. Cet [ɑ] n'évoluera pas vers [o] dans les langues, comme le grec et le latin, où la variante d'arrière de l'ancienne voyelle unique avait dû se différencier assez vite : le latin a *o* dans *potis*, originellement « le maître », « celui qui peut » (*pot-erat* « il pouvait », « il était maître »), restant bien distinct du *a* de *caput* « tête ». Au contraire, le sanskrit a, dans *kapucchala* « chignon », le même *a* que dans *patis* « époux ». Le germanique ne distingue pas non plus *a* de *o* : au latin *caput* correspond le vieil-anglais *hafud* avec *a*, tout comme le gotique *-faþs* « époux » qui correspond à *potis* (dans *brūþ-faþs* « celui qui est en puissance d'épouse »). Distinguent **a* de **o*, outre le latin et le grec, les langues italiques, le celtique, l'arménien et le tokharien; les confondent, outre l'indo-iranien et le germanique, le hittite, le baltique et le slave. Dans le cas du slave, en particulier, on peut se demander si l'indistinction n'est pas le résultat d'une évolution récente.

Indo-iranien mis à part, la distinction entre **e* d'une part et une voyelle alternante **o* ou **a* est, elle, générale. On n'est pas arrivé, à ce jour, à dégager dans quelles conditions s'est établie cette distinction. Il est très vraisemblable que la place de l'accent, à un certain point de l'évolution de la langue, a dû jouer un rôle.

La nature de l'accent, les conditions de son apparition, ses rapports avec la mélodie du discours changent au cours du temps, comme le font les phonèmes et les relations qu'ils entretiennent entre eux. Chacune des langues indo-européennes attestées a son propre système accentuel et la comparaison permet au mieux de supposer qu'à une certaine époque les voyelles ont été conservées sous l'accent et ont tendu à disparaître ailleurs. On parle de **degré plein** de la voyelle lorsqu'on la retrouve sous la forme de *e* ou de *o*, de **degré zéro** lorsque la voyelle a disparu, et de **degré réduit** lorsqu'elle a laissé une trace qui

a retrouvé, par la suite, avec un timbre particulier, un statut de voyelle pleine. Dans une syllabe inaccentuée on attend normalement zéro, c'est-à-dire chute de la voyelle. Mais lorsque, de ce zéro, pouvaient résulter des groupes de consonnes difficilement prononçables, on peut penser qu'on a utilisé une voyelle d'appui, un peu comme les Français utilisent un *e* dit « muet » pour alléger les groupes de consonnes dans *match(e) de football, arc(que) de triomphe, ours(e) blanc*, ou *Félix(e) Potin*.

La première idée qui vient à l'esprit pour expliquer l'alternance *e/o*, lorsqu'on ne projette plus les reconstructions sur un seul écran, mais qu'on essaie de différencier des stades successifs, est de supposer que, sous l'accent, la voyelle a pris un timbre antérieur, [æ], qu'elle est tombée ailleurs que sous l'accent, mais que là où sa chute aurait abouti à des groupes de consonnes imprononçables, elle s'est maintenue ou a été restituée avec son timbre plus profond [ɑ] ou [å].

Là où, dans la reconstruction traditionnelle, on parle de voyelles réduites, on doit avoir affaire à des phénomènes plus tardifs, particuliers à certaines langues ou groupes de langues, avec finalement des voyelles pleines restituées dont le timbre varie de langue à langue, *a* en latin et *u* en germanique, par exemple.

Au départ, [æ] et [å] n'étaient que deux variantes d'un même phonème dont la répartition dépendait de la place de l'accent. Mais la moindre modification du système accentuel pouvait aboutir à assurer, à l'un et à l'autre, leur indépendance phonologique, /æ/ et /å/ pouvant désormais figurer aussi bien sous l'accent qu'ailleurs. Dès lors, la porte était ouverte à toutes les extensions analogiques possibles. « Du pied » a dû, très anciennement, se dire **pedés*. On a, en latin, *pedis*, pour **pedes*, avec deux *e* ; en grec *podos*, avec deux *o* ; on attendrait en fait un **pdes* qu'on a vite rendu plus aisément prononçable, soit au moyen de la voyelle d'appui, d'où **podés*, soit par extension analogique de la voyelle attendue à l'accusatif **pedm̥*, donc **pedes*, ce qui donne la forme latine. Dans d'autres contextes, la désinence *-és* du génitif a pu perdre son accent et voir sa voyelle restituée sous la forme de *o*, d'où **podos*, comme en grec. Mais, bien entendu, dans cette veine, on peut expliquer à peu près n'importe quoi, ce qui semble, pour beaucoup, impliquer que l'explication n'en est pas une.

En désespoir de cause, il faudrait poser, pour des formes reconnues comme très anciennement établies dans la langue, une voyelle **e/o* qui a toutes chances d'avoir été au départ un /a/ avec des réalisations entre [æ] et [å]. On réserve **-e-* et **-o-* pour des spécialisations plus tardives, comme **-es* marque de pluriel alternant avec **-s*, mais n'alternant pas avec **-os*, comme *-e*, marque des vocatifs des radicaux à voyelle, comme **-om* désinence de génitif pluriel ou **-o-*, voyelle radicale du parfait, etc. Toutefois, on ne se résoud pas à employer constamment le

digraphe **e/o* qui alourdit les textes et ne facilite pas l'identification des formes qui ne sont plus lisibles à voix haute. On est tenté de noter simplement **e* en dépit de son ambiguïté.

Comme on le verra plus loin, on n'a jamais cessé de flirter avec l'hypothèse de la voyelle unique d'un stade très ancien de l'indo-européen. Inspirée à ses débuts par la structure du sanskrit, elle a aujourd'hui le soutien de l'analogie de certaines langues caucasiques où, au-delà de la variété des timbres vocaliques perçus, on peut poser, au moins dans le vocabulaire indigène, un phonème vocalique unique qui laissera au contexte le soin de décider de son timbre. Si, conventionnellement, nous notons cette voyelle comme ə, *pə* se prononcera [pø], le jeu des lèvres pour [p] entraînant l'arrondissement de la voyelle ; *tə* se prononcera [te] parce que la fermeture du [t] se réalise à peu près là où s'articule [e], et ainsi de suite.

Les voyelles longues.

L'autre remise en cause, qui aura finalement de profondes répercussions sur la façon de concevoir la reconstruction indo-européenne, porte sur le statut des voyelles longues. Parmi celles qui sont attestées dans les diverses langues et qu'on reconstruit pour l'indo-européen commun, on peut distinguer deux types. Tout d'abord celles qui alternent avec la voyelle **e/o* dont le timbre est le même, donc **ē* et **ō*. C'est, par exemple, le cas de la voyelle du nominatif latin *pēs*, en face de l'accusatif *pedem* et du génitif *pedis*, celui de l'équivalent grec *pōs* (écrit *poús*), en face de l'accusatif *póda* et du génitif *podós ;* ou encore celui du génitif pluriel grec du même mot, *podôn* en face de l'équivalent latin *pedum* pour **pedom*.

Ensuite, il y a les voyelles longues qui alternent avec ce qu'on appelle le *schwa*, c'est-à-dire la voyelle qui est *a* dans les langues européennes, mais *i* en indo-iranien. C'est ainsi que pour « se tenir debout », angl. *stand*, on a en sanskrit le participe *sthita-* en face de la base *sthā* et en latin *status* en face de *stā-re ;* pour « donner » on a en latin un *a* dans *datus* en face du *-ō-* de *dō-num* « don » ; pour « semer » le participe latin est *satus* avec *a* en face du *-ē-* de *sēuī* « j'ai semé ». Ceci veut dire qu'au participe, dans une syllabe à l'origine inaccentuée où l'on pourrait attendre une voyelle zéro, on peut poser *ə, mais que là où l'on attend une voyelle pleine, on a une longue de timbre particulier : *-ā-*, *-ō-* ou *-ē-*. On doit à Ferdinand de Saussure la suggestion que le *schwa*, *ə, se comporte ici comme une sonante, une semi-voyelle comme *i, y* par exemple : lorsqu'une racine comme **leikʷ-* « laisser » perd son accent, la voyelle (**e/o*) disparaît et la sonante *i* devient le sommet de la syllabe : en grec, par exemple, en face de *leípō* « je laisse » on a *élipon* « je laissai » ; de façon analogue, le

lat. *satus* « semé » se reconstruit comme **sɔto-s*, où *-ɔ- est ce qui reste, après transfert de l'accent sur une autre syllabe, d'une racine **seɔ-*, dans *sēuī* « j'ai semé »; la voyelle longue *-ē-* provient de la simplification de la « diphtongue » *-eɔ-*, de la même façon que la diphtongue *ei* du grec *leípo* s'est vite réduite à une voyelle longue [i :].

On relève d'autre part que, dans les bases où l'on reconstruit un **ā*, il y a normalement alternance avec le *schwa*. Il en va de même avec ceux des **ō* qui n'alternent pas avec **ē*. On est donc amené à penser que c'est le *schwa* qui est responsable, dans ces deux cas, du timbre de la voyelle. Il faudrait donc distinguer entre trois *schwa :* un, qu'on note $ə_1$, ne paraît pas affecter le timbre de la voyelle, qui va donc apparaître comme *ē* ou comme *ō* dans les mêmes conditions où l'on trouve, sans *schwa* suivant, *e* et *o ;* un autre qu'on note $ə_2$ aurait pour effet de « colorer » la voyelle en [*a*] ; un troisième qu'on note $ə_3$ la « colorerait » en [o], supprimant ainsi les conditions d'une alternance avec *e*.

Les « laryngales ».

Les langues sémitiques illustrent bien ce type d'action de certaines consonnes sur les voyelles voisines; comme ces consonnes y sont désignées, souvent à tort d'ailleurs, comme des « laryngales », ce terme a été vite employé pour nommer les trois différents *schwa*. Tout ce qui vient d'être exposé et ce qui va suivre forme ce qu'on appelle la « Théorie des laryngales ». Pendant longtemps, cette théorie a conservé le tour largement algébrique que lui avait donné Saussure. On y opérait avec des formules comme $*eə_2 = *\bar{a}$, sans chercher à savoir ce que pouvait être physiquement $ə_2$ et l'on continuait à utiliser, pour noter les « laryngales », le signe vocalique ə. Puis l'habitude s'est établie d'utiliser *H* accompagné des mêmes chiffres, sans que cela entraîne nécessairement une réflexion relative à la nature des sons en cause.

Parallèlement aux observations relatives aux rapports des voyelles longues et du *schwa*, on a relevé que le **a*, celui qui est *a* dans toutes les langues, y compris l'indo-iranien (à la différence du *schwa* qui semble, dans ce dernier, être représenté par *i*), n'apparaît guère qu'à l'initiale du mot et dans des formations familières du type **kaput*. On pense à ces piliers de la comparaison que sont le mot pour le champ, lat. *ager*, celui qui veut dire « conduire », lat. *ago*, la racine *ak-* en référence à l'acuité, l'acidité, etc. De façon analogue, c'est généralement à l'initiale des mots que l'on trouve des **o* qui n'alternent pas avec *e*, comme dans lat. *odor* « odeur », *octō* « huit », *ouis* « mouton », *os* « os ». Mais ici on peut toujours arguer que, par hasard, les formes alternantes en *e* ne sont pas attestées.

On a donc pensé que ces vocalismes avaient été déterminés par une « laryngale » colorante initiale disparue, donc H_2*egro-* > *agro* « champ », H_3*ewi-* > *owi-* « mouton », etc. Bien entendu, la « laryngale », ici *préposée* à la voyelle, n'a pas de pouvoir allongeant et la seule trace de son existence est la coloration.

Les « laryngales » du hittite.

Pendant une cinquantaine d'années, les comparatistes responsables n'ont pas osé opérer avec ce qu'on considérait comme des grandeurs algébriques parfaitement hypothétiques. Cela ne voulait pas dire qu'au moins certains d'entre eux n'en tenaient aucun compte dans leurs réflexions : on comprend mal les écrits de Meillet, par exemple, si l'on en fait complètement abstraction, mais on n'y trouve pas de *schwas* numérotés. C'est le témoignage du hittite qui a finalement convaincu certains chercheurs qu'on pouvait se risquer à en faire ouvertement état. Les transcriptions qu'on donne, en caractères latins, des textes hittites attestés en cunéiforme, présentent fréquemment le signe *ḫ*, souvent simplifié en *h*, qu'on pense représenter un phonème d'articulation analogue au *ach-Laut* allemand. Ce signe apparaît très fréquemment dans des mots où les esprits aventureux avaient postulé une « laryngale ». On a, par exemple, un verbe *newaḫḫ-* qui correspond exactement au latin *nouā-re* « innover » ; « os » se présente en hittite sous la forme *ḫaštai* qui correspond au H_3*est-* postulé, en termes de « laryngales », pour ce mot. On s'est alors avisé que l'arménien présentait fréquemment un *h-* dans les formes de ce type, dans *hot* « odeur », *haw* « grand-père », lat. *auos*, *haw* « oiseau », lat. *auis*. Mais il y a, en arménien, des cas de *h* initial tardif qui affectent la valeur du témoignage.

Le hittite ne confirme que partiellement la théorie. Pour autant que les rapprochements entre cette langue et les autres sont convaincants, les *ḫ* du hittite apparaissent bien là où on postulait des « laryngales », mais on postule aussi des « laryngales » là où il n'y en a pas de traces en hittite. On pose, par exemple, une « laryngale » pour expliquer le *a* initial de sanskrit *ánti*, grec *antí*, lat. *ante* « devant, avant » et on la trouve dans hit. *ḫanti*, de même sens. Mais celle qu'on pose à l'initiale de skrt *apa*, grec *apo*, lat. *ab* ne se trouve pas dans l'équivalent hittite *apa*. On a donc supposé qu'il n'y avait pas une seule « laryngale » qui colorait en [ɑ], mais deux, dont une seule a été conservée en hittite. C'est ici que cesse l'algèbre et qu'on va essayer de se représenter deux sons différents susceptibles, l'un et l'autre, d'attirer vers l'arrière les voyelles voisines. Celle qui est attestée aurait été une fricative vélaire, ou mieux, uvulaire (*ach* allemand, *jota* espagnole) ; celle qui serait tombée aurait été une articulation plus profonde de type pharyngal.

Il faudra donc poser deux H_2, une uvulaire et une pharyngale.

Comme le *a* et le *o* ne sont pas distingués en hittite, on n'y peut faire le départ entre les « laryngales » de coloration *o* et celles de coloration *a*. Là où d'autres langues réclament un timbre *o*, on se trouve dans la même situation que pour *a* : la « laryngale » attendue est attestée dans certains cas, *ḫastai* « os », par exemple, absente ailleurs, dans *pasi* « boire, avaler », en face de lat. *pō-tāre* « boire », de *peH_3*. Il faudra donc postuler deux « laryngales » de coloration *o*, dont l'une se maintient et l'autre disparaît en hittite, donc deux H_3, ici encore une uvulaire et une pharyngale, en laissant pour l'instant de côté le trait articulatoire qui devait distinguer les H_3 des H_2.

En ce qui concerne la « laryngale » H_1, à laquelle on n'attribue pas de pouvoir colorant et qui va donc laisser aux voyelles voisines leurs timbres alternants **e* et **o*, elle ne semble pas avoir laissé de trace en hittite dans *te-* « placer », grec *ti-thē-mi*, ou à l'initiale de *es-*, le verbe être et de *ed-* « manger », lat. *edō*, angl. *eat*. Il y a toutefois quelques formes où *ḫ* hittite apparaît au voisinage de *e* et qu'on rapproche de formes grecques ou latines en *ē*, par exemple, *meḫur*, interprété comme « point du temps », rapproché du lat. *mētior* « je mesure ». Ce qui est noté *ḫ*, ici, n'avait donc pas de pouvoir colorant. Il est remarquable que, lorsqu'il se manifeste entre voyelles, cet *ḫ* apparaît simple, alors qu'il est doublé lorsqu'il est en contact avec *a*. Il est donc vraisemblable que dans *meḫur*, *ḫ* note un autre phonème que *ḫḫ* dans *newaḫḫ-*. En général, en hittite, la graphie double entre voyelles correspond à une sourde dans les autres langues, la graphie simple à une sonore. Mais cela veut dire simplement qu'il y avait en hittite deux phonèmes distincts et pas forcément que la différence entre eux était, dans cette langue, de sourde à sonore. On peut donc penser que *-ḫḫ-* correspondait à une fricative vélaire ou pharyngale entraînant une poussée en arrière de la masse de la langue, d'où les *a* voisins, alors que le *-ḫ-* unique notait une glottale sans effet sur le comportement de la masse linguale. Dans une notation phonologique on aurait dans un cas /newax-/, dans l'autre /meʾur/ avec un /ʾ/ qui indique une articulation, non plus pharyngale qui impliquerait rétraction de la langue, mais glottale, c'est-à-dire au niveau de la trachée, là où la langue n'est plus en cause.

En résumé, la confrontation de la théorie « laryngale » avec les données hittites où les deux colorations ne sont pas distinguées, amènerait à poser le système suivant :

	attesté	non attesté
colorantes (fond de la bouche)	x	ħ
non colorantes (glottales)	ʾ	h

Aspiration et sonorisation.

Mais coloration et allongement de la voyelle précédente ne sont pas, selon la théorie, les seules traces laissées par les « laryngales » disparues. On leur attribue également la possibilité soit d'aspirer la consonne qui les précédait, soit de la sonoriser. L'aspiration est essentiellement attestée en sanskrit : soit la racine bien connue qui indique la station debout, celle d'anglais *stand*, all. *gestanden*, celle qu'on trouve dans le mot *sta-tion* lui-même et d'où dérive l'imparfait *était*, du latin *sta-bat*. On reconstruit **stā*, donc **steH₂*-, mais le sanskrit présente *sthā*- avec une forme réduite dans *sthiti*- « station debout » : à partir de **steH₂*- on attend *stā*, mais si la voyelle tombe, on a **stH₂* suivi d'une voyelle de soutien, *i* en sanskrit[1], devant le suffixe *-ti* qui suit ; **stH₂iti*- donne *sthiti* et le *-h* est étendu aux autres formes de la racine. Pour les autres langues, on peut envisager, soit que l'analogie a favorisé les formes sans aspiration, soit que cette aspiration a généralement disparu sans laisser de traces.

La sonorisation par « laryngale » suivante a, en revanche, laissé des traces un peu partout. Elle est nette dans les formes qui signifient « boire ». C'est à elle que le français doit le *b*- initial de *boire* en face du *p* de *potable* : on pose une racine *peH₃*, manifeste dans le lat. *pō-tāre* « boire ». A partir de cette même racine, avec le *p* initial redoublé et suivi d'un *i*, procédé fréquent pour marquer la continuité de l'action, on aura **pi-pH₃e*- (avec degré zéro de la voyelle entre **p* et **H₃*). **H₃* devait être ici sonore, prononcé donc avec des vibrations de la glotte, car il a sonorisé le *p* immédiatement précédent, comme l'atteste le sanskrit *pibati* « il boit » et son équivalent exact irl. *ibid* (avec *p*- initial régulièrement disparu). L'aberrance de la succession *p...b...*, dans cette forme à redoublement de la consonne initiale, a été éliminée à Rome au profit de *-b-* dans *bibit* « il boit », mais à Falérie, à dix lieues plus au nord, en faveur de *p*-, dans *pipafo* « je boirai ».

Comme la théorie « laryngale » a été longtemps ressentie comme passablement scandaleuse, ceux qui la défendaient avaient l'impression que le scandale ne pouvait que croître avec le nombre de « laryngales » postulées. Il fallait donc se montrer économe en la matière et en réduire l'effectif au strict minimum. Sur la foi des deux exemples précédents, on a décidé que coloration *a* allait de pair avec aspiration, coloration *o* avec sonorisation. Dans le même sens, nous l'avons vu, certains refusaient d'opérer avec une coloration *o* en faisant valoir qu'un *o* pouvait toujours s'expliquer comme l'alternant de *e*, même en l'absence de formes attestées avec ce *e*.

1. On retient ici la suggestion de T. Burrow, *The Sanskrit Language*, Londres, s. d., p. 105 et 106.

Or, tout tend à prouver que l'indo-européen commun n'était pas riche en voyelles distinctes et, dans ce cas, sur le modèle des langues attestées où la situation est identique, on peut s'attendre, en compensation, à un consonantisme très varié. Sans doute pourrait-on faire valoir la richesse du beau système sanskrit. Mais comme nous l'avons déjà entrevu et comme cela se confirmera ci-dessous, de ce système carré de cinq séries et cinq ordres, donc de 25 unités[2], il n'en restera, pour l'indo-européen reconstruit le plus ancien, que 13; si l'on y ajoute la sifflante et les sonantes, le total sera de 18 unités. Ceci, se combinant avec la voyelle unique, nous donne 18 syllabes distinctes du type consonne + voyelle, ce qui est ridiculement peu. Le français parisien contemporain en comporte 234, dont certaines, sans doute, ne sont pas attestées, mais qui sont toutes disponibles. On ne voit donc pas pourquoi, une fois qu'on accepte de postuler l'existence préhistorique de phonèmes non attestés sous forme de segments distincts, on refuserait d'aller au-delà de trois ou de quatre.

La conservation des distinctions.

Il y a, en faveur de l'expansion du système des « laryngales », un argument auquel ne seront sensibles que ceux qui ont relevé qu'au cours d'une évolution, un trait phonique ne se maintient que s'il permet à lui seul de distinguer les formes les unes des autres. C'est un fait établi que la pertinence distinctive se conserve à travers le temps sous une forme ou une autre, les traits redondants cédant finalement à l'usure. Si donc une « laryngale » s'est maintenue sous la forme d'une aspiration ou si elle a pu imposer sa sonorité à ses voisins, c'est que cette aspiration et ce voisement étaient *à eux seuls* distinctifs. En d'autres termes, si le H_3 de *pe*H_3- « boire » a pu sonoriser le *p*- après chute de la voyelle, c'est qu'il y avait dans la langue un autre H_3 qui, lui, n'était pas sonore.

Cela se vérifie un peu partout : en français, par exemple, sont physiquement sonores, c'est-à-dire accompagnés de vibrations de la glotte, les phonèmes /b v d z ž g m n ñ l r y/, mais cette sonorité n'est distinctive que pour /b v d z ž g/ qui s'opposent un par un à /p f t s š k/ ; elle ne l'est pas pour les six autres puisqu'il n'y a pas en français de phonèmes /m n ñ l r/ et /y/ sourds. Or, lorsque l'un des six premiers est précédé d'un phonème sourd, sa sonorité, pertinente, s'impose à ce dernier; dans *anecdote*, le /d/ sonore impose sa sonorité au /k/ précédent qui devient /g/. Mais dans *réclamer*, par exemple, la sonorité *non pertinente* du /l/ n'a aucun effet sur la sourdité du /k/ ; aucun Français ne tend à dire **réglamer* au lieu de *réclamer*. La sonorité *non*

2. En réalité 24, puisqu'il y manque une nasale vélaire.

pertinente de /n/ n'amène jamais les locuteurs à dire **shrabnel* au lieu de *shrapnel*, **Davné* au lieu de *Daphné*.

Tout cela amène à poser deux phonèmes H_3, un sonore et un sourd, ainsi que deux phonèmes H_2 dont l'un était sourd non par hasard, mais parce que cette sourdité, attestée plus tard sous la forme d'une « aspiration », le maintenait distinct de l'autre.

Il est intéressant de noter que ces phénomènes d'aspiration et de sonorisation ne paraissent être le fait que des « laryngales » colorantes. Ceci confirme les conclusions auxquelles on arrivait ci-dessus, à l'issue des observations relatives au hittite, selon lesquelles les non colorantes étaient des glottales qui, par nature, ne se combinent pas avec les vibrations glottales puisqu'elles sont elles-mêmes glottales, caractérisées l'une /ʔ/ par le verrouillage de la glotte, l'autre /h/ par son ouverture. Ces deux traits, verrouillage et ouverture, excluent normalement les vibrations du même organe.

On est donc amené à poser le système suivant où, puisqu'il ne s'agit plus seulement du hittite où *a* et *o* sont confondus, on distinguera les deux colorations différentes, la seconde marquée, en anticipant ce qui sera dit ci-dessous, par un exposant ^{w}.

	Non glottales Colorantes			
	[a] (H_2)		[o] (H_3)	
	sourdes	sonores	sourdes	sonores
Uvulaires	χ̣[3]	ʁ	χ̣^{w}	ʁ^{w}
Pharyngales	ḥ	ʕ	ḥ^{w}	ʕ^{w}
	Glottales Non colorantes (H_1)			
Fermeture	ʔ			
Frottement	h			

3. L'emploi de [χ] est ici conforme aux recommandations de l'Association phonétique internationale. En fait, il vaudrait mieux utiliser cette lettre grecque pour la notation de la *spirante* vélaire et réserver [x] à celle de la *fricative* uvulaire. Sur la différence entre spirante et fricative, cf. André MARTINET, *Eléments de linguistique générale*, Paris, 1960, § 2-24. Toutefois, lorsqu'on traite, comme ici, de processus évolutifs, il peut être utile d'avoir un signe qui ne précise pas la profondeur, vélaire ou uvulaire, du phonème en cause. Mais ce qui nous fait préférer ici [χ] à [x] est le fait que l'emploi du [x] pour noter la « laryngale » serait très troublant pour le lecteur lorsque nous traiterons, ci-dessous, du durcissement de certaines « laryngales » aboutissant à des formes latines comme *senex* ou *imperatrix* où le x a sa valeur traditionnelle.

Il faut bien préciser qu'il n'est pas possible d'offrir des correspondances convaincantes pour chacun des phonèmes ainsi reconstruits. Il se trouve que dans les quelques cas où nous sommes amenés à postuler une sonorisation, celle-ci va de pair avec une coloration *o*, si bien que le χ^w et le $\d{h}^w$ du tableau pourraient sembler n'être là que par goût de la symétrie. La sonorisation des consonnes voisines réclame sans doute qu'elle ait été pertinente dans la « laryngale », donc qu'il y ait eu, dans le système, une « laryngale » sourde du même type en face de la « laryngale » sonore. Mais il est tout à fait possible qu'une « laryngale » sourde, sans partenaire sonore, se soit affaiblie en un [h] qui a ultérieurement aspiré la consonne précédente. Donc, ce qui vaut pour la sonorisation ne vaut pas pour l'aspiration. C'est l'expérience que nous avons aujourd'hui des systèmes phonologiques les plus divers qui nous amène à penser que si l'opposition de sourdes à sonores existait dans le système des fricatives que postule la théorie « laryngale », elle devait s'étendre à l'ensemble de ce système. On ne voit pas pourquoi une coloration *o* devrait aller de pair avec la sonorité et une coloration *a* avec son absence. Il faut donc interpréter la présence d'un phonème sur le tableau non comme l'affirmation de son existence à un stade de l'évolution de la langue, mais comme l'indication qu'un tel type phonique peut avoir figuré parmi ses phonèmes et qu'il ne faudrait pas hésiter à le poser si des faits nouveaux ou de nouvelles interprétations de faits connus le réclamaient.

La coloration a.

Il nous faut revenir sur ce que nous avons désigné comme des colorations. C'est, on l'a vu, l'expérience des langues sémitiques, et en particulier de l'arabe, qui permet surtout de comprendre comment le timbre de la voyelle peut être affecté par la consonne ou les consonnes voisines. La chose se retrouve ailleurs. En anglais, par exemple, la langue creusée en cuiller que réclame l'articulation des /r/ et /l/ de cette langue, entraîne des prononciations du /ɪ/ qui se rapprochent beaucoup de [u], par exemple dans *pretty* /'prɪti/ qui devient parfois ['pruti], dans *silk* /sɪlk/ et le nom *Wilson* /'wɪlsn̥/ qui peuvent être perçus respectivement comme [suk] et ['wusn̥]. En français où, pendant des siècles, le *-ss-* de *lasse* s'est distingué du *c* de *lace* par le relèvement vers le palais de la pointe de la langue, la prononciation parisienne, où *lace* [las] est bien distinct de *lasse* [lɑs], porte jusqu'à ce jour témoignage de l'action de la consonne sur l'articulation de la voyelle[4].

4. Voir André MARTINET, *Economie des changements phonétiques*, Berne, 1955, chap. 9, § 9-15, et Martin JOOS, The medieval sibilants, *Language 28*, p. 222 et s.

En arabe, cette action se manifeste à grande échelle dans ce qu'on appelle l' « emphase ». Normalement un *ā* arabe s'articule plutôt vers l'avant de la bouche. La rivière, ou la vallée qu'elle a creusée, *wād(i)*, est attestée avec un *a* dans le nom de fleuves espagnols comme le *Guadalquivir* (littéralement « rivière-la grande »), mais elle se réalise en fait aujourd'hui comme [wæ:d] que les Français reproduisent comme *oued*. On notera aussi l'argotique *(pas) bezef* « pas beaucoup », pour l'arabe *bezzāf*. Si toutefois le /ā/ de l'arabe se trouve en contact avec une consonne dite « emphatique », il prend un timbre très profond, celui de *lasse* dans les prononciations dites « Marie-Chantal » du XVI[e] arrondissement parisien. L' « emphase » est, en fait, la poussée, vers l'arrière de la cavité buccale, de la masse de la langue qui permet de distinguer *tāb* [tæ:b] « il s'est repenti » de *ṭāḅ* [ṭɑ:ḅ] « il a cuit » (le point souscrit note ici l' « emphase »). Pour un Arabe, qui sait réaliser l' « emphase », les voyelles, bien que physiquement différentes, sont ici les mêmes, la différence première et décisive se trouvant dans les consonnes. L'étranger, qui ne connaît pas ce phénomène, perçoit la différence entre les voyelles. Ce qu'on appelle un *ṭ* « emphatique » ajoute à l'articulation du [t] cette poussée de la langue vers l'arrière.

Cependant cette poussée, au lieu d'être un trait qui s'ajoute aux autres traits articulatoires du phonème — dans le cas de /ṭ/, l'articulation de la pointe de la langue et celle de la glotte —, peut représenter l'essentiel de l'articulation de la consonne. On a, dans ce cas, une uvulaire ou, avec une poussée vers l'arrière plus accentuée, une pharyngale. Les voyelles voisines seront alors affectées de la même façon que lorsque l' « emphase » est un trait ajouté à l'articulation de la consonne. C'est, en fait, ce que nous avons désigné comme une coloration.

On n'a pas de raison de supposer, à aucun stade envisageable de l'indo-européen, un phénomène, analogue à l' « emphase », s'ajoutant à une autre articulation consonantique. Mais ce qu'on y postule, ce sont des phonèmes d'articulation profonde qui se sont affaiblis au cours du temps pour finalement disparaître dans toutes les formes de l'indo-européen parlé à ce jour. L'affaiblissement étant graduel, l'influence de l'articulation consonantique sur celle de la voyelle voisine en est venue à représenter la consonne elle-même. Si nous notons la consonne en cause comme /χ/ et la voyelle ouverte de la langue comme /a/, la succession /χa/ se prononcera en fait [χɑ] avec un [ɑ] profond. Si, dans la même langue, il existe une glottale /h/ sans effet sur la voyelle adjacente, /ha/ se prononcera [ha] avec un [a] moyen. Si maintenant le /χ/ s'affaiblit pour finalement se confondre avec [h], le frottement audible étant finalement celui qui se réalise au niveau de la glotte, l'ancienne distinction entre /χa/ et /ha/ se

maintiendra sous la forme de /hɑ/ et /ha/, puis, si le /h/ s'affaiblit à son tour, sous celle de /ɑ/ et /a/. Mais dès que /ɑ/ et /a/ deviennent deux unités distinctes, deux *choix* différents, ils vont tendre à accentuer leurs différences, /a/ tendant vers [ɛ] et [e]. C'est ce qui explique qu'on ait, en latin, par exemple, *agō* (< *$*H_2eg\bar{o}$*) « je pousse » distinct de *ego* « moi » (< *$*H_1eg\bar{o}$*).

La coloration o.

La coloration *o* suppose également une poussée de la langue vers l'arrière, mais, en plus, un arrondissement concomitant des lèvres. Qu'un tel arrondissement puisse se combiner avec une articulation dans une autre partie de la bouche est bien connu, et on le pose traditionnellement comme un trait de l'indo-européen commun. C'est lui qui se manifeste dans les labiovélaires du type $*k^w$ attestées jusque dans le latin *quattuor* et, sous la forme d'éléments successifs, dans l'italien *quattro* et l'espagnol *cuatro*. On peut donc supposer qu'il existait, non seulement en combinaison avec une occlusive comme [k], mais aussi avec une fricative comme [χ], de même qu'on a, en espagnol, *Juan* [χwan] à côté de *cuanto* [kwanto], même si, dans ce cas, les combinaisons [χw] et [kw] sont moins intimes que celles qu'on note $*k^w$ et $*\chi^w$ [5]. On suppose donc que la forme ancienne du mot qui désigne la branche, grec *ózos* (<*osdos*), all. *Ast*, hit. *hasd-wer*, est $*\chi^w$*e-sdo-*. Il semble que la formation de ce mot soit parallèle à celle qui a donné le français *nid*, lat. *nīdus*, angl. *nest* de **ni-sdo* : le second élément est, de part et d'autre, la racine *sed* « s'asseoir, se poser », au degré vocalique zéro ; le premier est, pour *nid*, l'élément adverbial *ni-*, bien identifié, qui marque un mouvement vers le bas ; pour la branche, $*\chi^w$*e-sdo-*, c'est une autre particule marquant une installation moins permanente, celle qu'on qualifie encore aujourd'hui au moyen de l'expression « comme l'oiseau sur la branche ».

A cette interprétation phonologique [χ^w] du H_3 algébrique, certains ont pu objecter que l'on s'attendrait, dans ce cas, à ce que l'élément marqué par [^{w}] se manifeste finalement comme [w], comme on le voit dans l'espagnol *cuatro* avec *cu-* pour $*k^w$, et non pas, comme le postule la théorie laryngaliste, par un arrondissement de [ɑ] en [å] de la voyelle suivante. Ceux-là oublient que le type d'évolution du $*k^w$ initial du mot pour « quatre » jusqu'au /kw/ de l'espagnol *cuatro* est rare en comparaison de toutes les évolutions attestées vers *k*, dans les langues *satem*, vers *k* ou *p* en celtique, *p* puis *f* en germanique, *t-* en grec devant voyelle d'avant. Il est constant que $*k^w$ évolue vers *k* devant voyelle arrondie comme en lat. *cūr*, de $*k^w\bar{o}r$; lorsque $*k^w$ est, en germanique,

5. Voir André MARTINET, *Economie...*, chap. 8.

devenu h^w, l'arrondissement a disparu devant *ō* suivant comme le montre l'anglais *how* (< *hū* < h^w*ō*). Il faut bien comprendre que ce trait de la consonne n'a aucune raison de se maintenir comme segment distinct une fois qu'il s'est manifesté dans la qualité de la voyelle suivante.

Si, au lieu de précéder la voyelle, *χ^w, ou telle autre fricative labiovélarisée, la suit, il faut distinguer entre ce qui se passe si /$e\chi^w$/ appartient à la même syllabe, à la fin du mot, par exemple, ou devant consonne de la syllabe suivante. Dans ce cas *χ^w disparaît en allongeant la voyelle précédente, et l'arrondissement labial se manifeste sur cette voyelle. Le produit de cet amalgame est un [å:] qui garde tout le pouvoir distinctif de l'ancienne combinaison [$e\chi^w$] si, comme on peut le penser, cette voyelle longue se prononce sur un autre ton que les longues de même timbre résultant d'autres processus (cf. ci-dessus, p. 74). Nous notons désormais simplement *ō* pour la longue et *o* pour la brève correspondante.

Si maintenant /χ^w/ se trouve entre deux voyelles, il fait partie de la deuxième syllabe, la coupe étant *-*e*-χ^w*e*-. Il disparaît alors sans allonger la voyelle précédente. Mais on va mettre à profit l'arrondissement labial pour éviter l'hiatus et la seconde syllabe va apparaître comme *-we-*. Cet élément labial s'étant manifesté sous la forme de [w], il ne sera plus disponible pour changer le timbre de la voyelle précédente. Mais la poussée de la langue vers l'arrière qui ne se manifeste pas dans *-we-* se maintiendra sur cette voyelle qui conservera la coloration *a* correspondant à cette poussée. Le résultat sera donc *-a-we-* (ou *-a-wo-* si, dans cette position, le *e* se trouve alterner avec *o*). Ceci est d'autant plus vraisemblable que cette seconde voyelle a des chances de faire partie d'une désinence commençant par une voyelle. Mais, naturellement, cette désinence va alterner avec d'autres commençant par une consonne. Le radical va donc se terminer, selon les cas, en *-aw-* ou en *-ō*.

Il en résultera des extensions analogiques aboutissant soit à l'établissement du timbre *o* dans la première syllabe, d'où *-owo-*, soit à l'allongement du *a*, d'où *-āwo-*. C'est ce qu'illustre l'ordinal formé à partir du cardinal « huit ». Celui-ci peut se reconstruire comme *H_3ekteH_3 où il semble que *-e*H_3 soit la marque d'un duel, le radical ayant désigné la main sans le pouce, donc quatre doigts ; *huit* est donc deux fois quatre doigts. L'ordinal est généralement formé par l'addition d'un *-o-* au radical du cardinal. On posera, phonologiquement, pour le deuxième H_3, pour des raisons qui apparaîtront ci-après, non pas /χ^w/, mais la sonore /γ^w/. La forme reconstruite pourra donc être *$H_3ekte\gamma^w o$, qui évoluera régulièrement en *oktawo-*. Les formes attestées en grec et en latin sont, en fait, respectivement, *ogdowo-* et *octāuo-*. La voyelle *-o-* devant *-w-* en grec est brève, comme

on peut l'attendre, mais son timbre est dû à l'analogie de celui du cardinal *oktṓ*. Le -ā- du latin a le timbre attendu, mais sa longueur est due à l'analogie du *-ō* long du cardinal. Le *-gd-* sonore du grec s'explique à partir d'une forme **χ^{w}ekʁ^{w}o-* avec un groupe de trois consonnes successives, inattendu dans le vocabulaire général, mais attesté et explicable dans ce domaine linguistique un peu particulier que représentent les cardinaux.

Les cardinaux.

Les cardinaux ne s'apprennent pas comme le reste du vocabulaire dans des contextes suggérés par l'expérience à transmettre, mais en comptant : ... *cinq, six, sept, huit*, etc. C'est pourquoi l'analogie joue beaucoup entre eux et très fréquemment par anticipation. C'est par anticipation des [t] de *trente, quarante*, etc., que *vingt-quatre*, par exemple, se prononce comme s'il était écrit *vingte-quatre*, que ce soit à Paris où l'on entend [vε̃tkatʁ] ou dans le Midi où la forme est [vεntökatrö]. C'est par anticipation de la sonorité d'*ogdowo-* qu'on expliquera celle de *hébdomos* « septième », en face de *heptá* « sept ». « Septième » est, avec une sonorité de même origine, en vieux-slave, *sedmŭ* qui suppose également un groupe de trois consonnes dans **septmo-* > **sebdmo-*. Les ordinaux sont souvent le produit d'une dérivation très vivante, chaque locuteur étant susceptible d'en fabriquer un, dès qu'il en a besoin, avec des résultats phonétiques souvent inattendus. Le groupe [-tþ] de l'anglais *eighth*, que ne reproduit même pas la graphie, est parfaitement isolé dans la langue. En français, on entend parfois *quatrième* en deux syllabes, [katʁjεm], au départ, sans doute, dans la série des dissyllabes *deuxième, troisième... cinquième*, chez des sujets qui prononcent pourtant *ouvrier* en trois syllabes, [uvʁije], selon la norme qui veut qu'après un groupe consonne + *r* et devant voyelle, *i* se prononce comme s'il était écrit *-ill-*, donc **ouvriller*.

La « laryngale » labiovélaire.

L'interprétation de H_3 comme labialisée sonore ou sourde, avec, peut-être, deux degrés de profondeur articulatoire : χ^w, ʁ^w uvulaires, h^w et ʕ^w pharyngaux, a retenu l'attention, car elle offrait une explication séduisante du [w] caractéristique des parfaits latins comme *gnōu-ī* « j'ai connu », d'un radical **gneH^w-* suivi immédiatement des désinences caractéristiques. La résistance qu'elle rencontre auprès de certains est due à l'incompréhension du principe fonctionnel selon lequel un trait phonique se maintient dans la mesure où il assume une valeur distinctive. Il faut aussi tenir compte de réticences devant

l'utilisation à grande échelle de l'analogie, dont on a l'impression qu'elle peut tout expliquer, et que, par conséquent, elle n'explique rien. Mais lorsqu'on considère un problème spécifique, comme celui du parfait indo-européen pour lequel les langues se sont toutes données des marques particulières, il faut nécessairement supposer que chaque langue a retenu un (ou parfois plus d'un) des traits qui se trouvaient caractériser par hasard certains parfaits pour l'étendre aux autres par analogie.

Bien avant qu'on parle de « laryngales », les latinistes avaient émis l'opinion que le parfait en [w] s'était étendu par analogie à partir de radicaux à voyelle longue finale. Comment écarter l'analogie lorsqu'on constate que ces verbes à voyelle longue présentent tous en anglais des formes à *-w* final, non plus au parfait, mais à tous les temps, en face d'équivalents allemands qui ignorent radicalement le *-w* : en face de lat. *flāre* « souffler », parfait *flā-u-*, on a anglais *blow*, mais all. *blähen*. De même *sew-sähen, mow-mähen, glow-glühen*, etc., où il ne faudrait pas croire que le *-h-* des formes allemandes est l'équivalent phonétique régulier du *-w* anglais.

Il est remarquable que partout où l'on reconstruisait traditionnellement un **ō* long, on le trouvait alternant avec des formes à *-w-* souvent précédé d'une voyelle brève de timbre *a*, ce que laisse attendre une évolution fonctionnelle de la phonologie de la langue à partir d'une labiovélarisée.

Pour illustrer différentes implications de la théorie des « laryngales » labiovélarisées, on retiendra ici le cas de **prō* représenté en français par la préposition *pour* et le préfixe *pro-* « favorable à ». La valeur première est « vers l'avant ». Au départ il y a **per* « à travers », auquel s'ajoute un suffixe *-eH*^{w} que nous nous hasardons à expliciter phonologiquement comme /χ^{w}/ pour le distinguer du /ʁ^{w}/ qu'on postule pour le suffixe de duel, celui de **oktō* « huit ». Le composé **pr-e*χ^{w}, avec degré zéro du premier élément, marquerait un mouvement à travers et au-delà. Le **prō* que l'on attend, dans les langues comme le grec ou le latin, où *ō* est distinct de *ā*, présente aussi un *o* bref, mais l'abrègement en finale est fréquemment attesté pour les différentes voyelles. A partir de cette base, la dérivation la plus simple est celle qui est réalisée au moyen de la voyelle *-o*. Les sens vont être « celui qui va en avant », « premier », « qui s'impose », « correct », « juste », « vrai ». A date ancienne, seule la voyelle suffixale se maintient. On a donc **pr*χ^{w}*-o-* dans lequel *-r-* se syllabise, ce qui donne régulièrement *pūrva* en sanskrit, *prŭvŭ* en vieux slave (russe *pervo-* comme premier élément de composé) « premier ».

Avec maintien de la voyelle du radical, on a **pre*χ^{w}*-o-*, d'où **prawo-* donnant en germanique *frawo-* qui désigne le premier dans la société, c'est-à-dire le seigneur. Seul le féminin a survécu en allemand sous la

forme de *Frau* « dame », évolution parallèle à celle du français où, de *dominus, domina* « maître, maîtresse », seul le féminin *dame* demeure, le masculin restant attesté dans *vidame* et des noms propres comme *Dampierre, Dammartin,* variantes de *Saint-Pierre, Saint-Martin.* Ces parallélismes évolutifs franco-allemands ne sont sans doute pas dus au hasard, mais, comme les genres différents du mot *See* masculin avec le sens de *(le) lac,* féminin avec celui de *(la) mer* (cf. ci-dessus, p. 43), ils sont attribuables au bilinguisme des temps carolingiens.

Du même **prey̆ʷo-*, avec allongement analogique de la voyelle, on a **prāwo-* qui est le russe *prav-yj* « droit, juste », d'où *pravda* « la vérité ». Curieusement, le correspondant formel latin *prāuos* a pris le sens du composé **dē-prāuos* « dépravé » (cf. *dēprāuatio* « dépravation »). Faut-il supposer au départ des emplois ironiques du type « C'est du joli ! » ? En grec, « premier » se dit *prôtos* à Athènes, *prātos* en béotien. Le suffixe *-to* est celui qu'on a dans les premiers ordinaux (sauf « deuxième ») en grec comme en latin ; dans cette dernière langue, on a *quartus, quintus, sextus.* En grec, il s'agit de sa variante *-ato* au suffixe (cf. *húpatos* « suprême ») ajouté à **prow-* ou **prōw-* à Athènes, à **praw-* ou **prāw-* en béotien.

La marque régulière du locatif *-i* se rencontre dans des adverbes à valeur temporelle comme grec *perusi, peruti* « l'année passée » de *per* + la forme à voyelle zéro *ut-* d'un radical dérivé de *wet-* « année ». Ce même *-i* de locatif, ajouté à **preχʷ*, a donné **prawi* qui, avec chute du *w* intervocalique, a donné le latin *prae* « avant, devant ». On a eu, par analogie, **prōwi,* avec le sens « de bonne heure », attesté dans le grec *prōí* et l'all. *früh.*

Le fr. *proue,* avant du navire, dérive, par l'intermédiaire du latin *prōra,* du grec *prō(i)rā,* de même sens, formé à partir de **preχʷ* au moyen du suffixe *-riā.*

Martin S. Ruipérez signale[6] le parallélisme des formes latines en *-tiō, -tiōnis* et en *-tīuus* (par exemple *actiō* « action » et *actīuus* « actif »). Il dérive ces formes du suffixe d'abstrait *-ti-,* bien connu, suivi de *-eH₃-*. La base suffixale **-tieH₃-* ainsi formée se réduit à **-tiH₃-* devant le suffixe adjectival *-o-*, d'où *-tiwo-*, et, avec un *ī* long analogique, *-tīwō*, qui donne le latin *-tīuus.* A la base suffixale, le substantif ajoute un « élargissement » *-n-* bien connu ailleurs. Au nominatif, **-tieH₃n-* aboutit en latin à *-tiō,* en osque à *-tiuf* (de *-tiōns*), produits d'évolutions phonétiques et de restitutions analogiques. Au génitif, on attend une base à voyelle zéro devant une désinence *-es,* donc **-tiH₃n* qui aboutit au *-tīn-* qu'on pose pour l'osque. La forme latine *-tiōnis* est naturellement analogique du nominatif.

6. Dans *Bivium, Homenaje a Manuel Cecilio Díaz y Díaz*, Madrid, 1983, p. 275-277.

*L'élargissement en -*k*-.*

Les comparatistes ont très vite relevé qu'à côté des suffixes proprement dits, comportant au départ une voyelle qui, naturellement, peut se maintenir ou disparaître, il existait des ajouts consonantiques auxquels on ne savait trop quelle valeur attribuer. On les désigne comme des « élargissements », ce qui ne fait que camoufler la difficulté qu'il y a à les identifier sémantiquement. Le *-w-* issu de *$*H_3$* permet de comprendre comment une consonne qui alterne avec zéro à la fin d'un radical peut n'avoir effectivement, à l'origine, aucune valeur. Elle peut certes en acquérir une ou plusieurs, ultérieurement, par expansion analogique, comme on le constate dans le cas du *-u-* des parfaits latins et dans certaines désignations de couleur, où le *-w-* peut dériver d'une base **bhleH*$_3$*-* qui a donné en latin, avec le suffixe *-o-*, *flāu-os* et avec le suffixe *-ro-*, *flō-rus* « jaune, blond ».

Un « élargissement » bizarre est le *-k-* qui se présente en latin, à la finale du radical d'un mot comme *imperātrīx* « impératrice ». Qu'en face du suffixe *-tor* du masculin *imperātor* on trouve, avec chute de la voyelle entre *-t-* et *-r-*, une forme féminine munie d'un *-ī-*, cela a toujours paru naturel, car les désignations d'êtres féminins en *-ī-* sont constantes en sanskrit et fréquentes ailleurs. Mais qu'ajoute le *-k-* ? On sait qu'un *ī* long qui se trouve alterner avec *-ia*, comme celui-ci, suppose une « laryngale », et l'on peut se poser la question de savoir dans quel contexte une « laryngale », c'est-à-dire une fricative dorsale profonde, a pu se durcir en une occlusive de même lieu d'articulation.

La première suggestion pour poser des laryngales à la source de certains *k* attestés vient d'Edward Sapir qui a suggéré que *-ka*, caractéristique régulière du parfait grec, représentait la succession de deux « laryngales »[7]. La première se trouvait à la fin de radicaux considérés traditionnellement comme se terminant en voyelle longue, c'est-à-dire voyelle + « laryngale », ceux-là même qui ont reçu tout d'abord, en latin, la marque *-w-* du parfait. La deuxième était la consonne initiale de la désinence *-H_2e* de première personne au parfait. Soit, par exemple, le radical *deH*$_3$*-* avec valeur de « donner » ; suivi de *-He*, on a *deH*$_3$*-He* où les deux laryngales vont s'amalgamer en *-k-*. On attend en grec **-doka*. On a, en fait, *dedōka*, avec allongement analogique du *o* et le redoublement de la consonne initiale, régulier dans les parfaits. Sturtevant[8] a ultérieurement cherché d'autres exemples de *k* issus de la rencontre de deux « laryngales », mais trop souvent en posant gratuitement certaines « laryngales » pour immé-

7. Cf. Edgar H. Sturtevant, *The Indo-Hittite Laryngeals*, Baltimore, 1942, p. 19.
8. *Ibid.*, p. 87-89.

diatement les faire disparaître dans l'amalgame [k]. Cet abus ne doit pas faire rejeter la suggestion de Sapir et il y a probablement d'autres cas où $*H + H$ a abouti à *k*.

Il y a toutefois, semble-t-il, une source beaucoup plus productive de [k] à partir des laryngales. C'est le durcissement des fricatives dorsales en occlusives devant un *-s* suivant qui est très bien attesté en germanique[9]. La graphie allemande témoigne encore aujourd'hui de l'ancienne valeur de fricative du premier élément de ce qui est devenu un groupe /ks/ : all. *Fuchs* « renard » aujourd'hui /fuks/ porte témoignage d'une ancienne prononciation /fuχs/ où /χ/ était le produit normal de l'évolution en germanique d'un plus ancien **k*.

Le cas le plus frappant est celui du latin *senex* où le /-k-/, impliqué dans *-x*, n'existe qu'au nominatif devant la désinence *-s ; senex* reflète régulièrement **senaks* qui, selon la suggestion qui précède, serait le produit normal de */$seneH_2$ + s/ avec, à l'origine, l'accent sur /-neH_2/. Avec l'accent sur la désinence *-es* du génitif, on aura chute de la voyelle entre *-n-* et -H_2, donc *$senH_2$-es*. La voyelle *-é-* en contact avec -H_2- devrait passer à *-a-*, mais c'est la désinence *-es*, normale, qui est rétablie par analogie. On obtient donc avec chute régulière de χ et élimination totale de toutes les implications possibles de la « laryngale », un génitif **senes* qui va normalement évoluer vers *senis*, forme latine attestée. Si à *$seneH_2$- on ajoute le suffixe *-tu-*, on obtient *$seneH_2$-tu- qui désigne le conseil des vieillards, lequel évolue régulièrement en *senātu-* « sénat ». Dans ce mot *senex*, très fréquent, ce qui implique que toutes les formes sont très tôt maîtrisées par l'enfant, aucune analogie n'a joué, ce qui donne des paradigmes de flexion et de dérivation parfaitement irréguliers et, de ce fait, extrêmement informatifs pour le comparatiste.

Le *s* qu'on peut rendre responsable du durcissement de la laryngale est, outre la marque du nominatif, celui de certains affixes verbaux qui ont donné, dans les langues attestées, des formes connues comme des aoristes, ceux notamment qui sont à la source de certains parfaits latins. Une forme comme *uīxit* « il a vécu », en face du présent *uīuit*, suggère immédiatement un radical terminé en -H_3 où la laryngale s'est durcie en [k] devant le *-s-* dit d'aoriste et où elle a laissé un [w] devant voyelle suivante. On pose *g^wiH_3, forme à vocalisme zéro d'une racine *g^weiH_3 qui, devant un suffixe en *t*, donne *$g^w\bar{\imath}$-, d'où le latin *uīta* « vie » et le russe *žit'* « vivre ». Devant voyelle suivante, on a *$g^w\bar{\imath}w$-, avec un *ī* long analogique, d'où le latin *uīuo* « je vis » et le russe *živ'ot* « il vit ». Avec le suffixe *-o*, on a, avec *ī* long analogique, lat. *uīuo-s* « vif », skrt. *jīva-*, v. sl. *živu*, et avec l'*i* bref régulier, grec *bios*

9. Cf. André MARTINET, Le couple *senex-senātus* et le « suffixe » *-k-*, *BSL* 51, p. 42-56, repris dans *Evolution des langues et reconstruction*, Paris, 1975, p. 146-168.

(<**biwos*) et got. *qiwa-*. Formé directement sur la racine, on aura, à partir d'un nominatif g^wiH_3-*s*, une forme g^w*ik-* alternant avec g^w*iw-* qui peuvent expliquer, par contamination mutuelle, le vieil anglais *cwicu*, d'où angl. *quick*.

Lorsqu'on cherche à retrouver des *-k* d'allure suffixale qui pourraient provenir d'une « laryngale » suivie de *-s* on s'aperçoit qu'à date ancienne, c'est-à-dire dans des termes ou des formations attestés largement et qui ne sont pas susceptibles d'être des innovations locales, de tels *-k-* n'existent qu'après des *a*, des *i*, des *u* longs ou brefs, c'est-à-dire là où l'on attend une « laryngale » qui explique le timbre *a* ou l'alternance de *i*, *u* brefs avec la longue correspondante. Il y a d'abord, bien entendu, des cas où le traitement de *k* est celui d'une laryngale qui fait partie de la racine même, comme dans le cas du mot pour « langue », russe *iazyk*, celui pour « poisson », baltique *zuk*, arménien *jukn*, et, nous venons de le voir, le cas du mot pour « vivre », « vif ». Mais les formations les plus fréquentes sont celles où l'on a, au départ, un suffixe en $-eH_2$ qui, en finale, va normalement aboutir à *-ā* ; il est souvent précédé d'un *-y-*, d'où $-yeH_2$ qui apparaîtra le plus souvent sous la forme réduite $-iH_2$-, donc *-ī* en finale.

Parmi les formes en *-ā*, il faut sans doute mettre à part celles où *-ā* désigne des féminins, qu'il s'agisse du sexe — mais là les formes sont plutôt récentes — ou du genre où elles sont un peu plus anciennes. Ces noms féminins en *-ā* ne prennent jamais le *-s* de nominatif, la raison en étant vraisemblablement qu'elles tiennent leur *-ā* d'un pronom démonstratif *asyntaxique*, **sā* « celle-là », comme nous le verrons ci-dessous, p. 189. Les masculins en *-ā* (< $-eH_2$) sont tout autre chose. Le suffixe y paraît avoir la valeur de « celui qui... » : le lat. *scriba* « le scribe » désigne celui qui a affaire à l'écriture, *agricola* « celui qui a affaire à la culture des champs », le russe *voevoda* « celui qui conduit les soldats ». Le nominatif, en latin, apparaît sans *-s*, ce qui est probablement dû à une chute régulière, à un certain temps de l'évolution, de *-s* après la « laryngale ». En grec, au contraire, on a un *-s*, réintroduit sans doute à date relativement tardive, par exemple dans *polítēs* « citoyen » (<**politā-s*), c'est-à-dire l'individu dans ses rapports avec la cité (*polis*).

Lorsqu'une forme linguistique résulte de la combinaison d'un radical et d'une désinence, il est assez normal que celui qui parle ne fasse que reproduire un ensemble qu'il a entendu précédemment ; ce sera nécessairement le cas si la forme est irrégulière. Mais il peut se faire que le locuteur combine de son propre chef le radical et la désinence. Si dans ce cas la forme socialement admise est irrégulière, il commettra une « faute » qui sera sanctionnée, mais qui pourra aussi être imitée par d'autres et se fixer finalement comme la forme correcte. Soit la forme $*sen\text{-}eH_2$-*s* restituée ci-dessus comme le nominatif de

$seneH_2$ « vieillard ». A un certain point de l'évolution de la langue, les usagers vont durcir *-χs-* en *-ks-*. On entendra donc [senaks]. Si la forme est fréquente, on la reproduira telle quelle. C'est ce qui s'est produit au cours de la préhistoire du latin. Mais, moins fréquente, elle aurait pu se voir remplacer, à chaque génération, par une combinaison du radical, tel qu'il se dégage des différentes formes de la flexion, et de la désinence, telle qu'elle se manifeste en général. Ici le résultat serait de remplacer [senaks] par [*$seneH_2$-s]. Mais, à ce point du temps, le [H_2] peut s'être un peu affaibli, la tendance étant à l'assourdir en fixant, comme précédemment le timbre [a], mais en transférant sa durée à la voyelle, le résultat final étant [senās]. A telle autre époque, ou dans tel autre usage ou dialecte de l'indo-européen, le groupe *H_2s*, jugé imprononçable, peut avoir tendu à perdre son [s], le résultat final étant [senā]. On peut donc, dans un cas de ce genre, s'attendre à trouver, comme produit d'un ancien *-eH_2-s*, soit 1. [-aks], soit 2. [-āks] par analogie d'un cas où la désinence est consonantique, mais autre que *-s*, soit 3. [-ās], le cas des masculins en *-ā-* du grec, soit 4. [-ā], ce qu'on a normalement en latin. Toutefois, dans cette langue, le produit 2. [-āks] est attesté dans une série d'adjectifs du type *audāx* « audacieux », *mordāx* « qui mord », etc. Le grec a des formes isolées de même type comme *thṓrāx* « cuirasse », *múrmēks* « fourmi » (< *mur-māks*). Partout, ici, le *k* s'est imposé par analogie à tous les cas. On a donc, au génitif, par exemple, *audācis*, *thṓrākos*, etc. Selon la tendance millénaire à remplacer les radicaux à consonne finale par des formes où cette même consonne est suivie de *-o-*, on aura des désignations d'individus en *-āko-*, fréquentes en slave, par exemple dans le nom *Novak*, qui a dû désigner le nouveau venu dans la localité. C'est probablement le même élément, en celtique, qui est à la source des noms des toponymes en *-ac* du Midi de la France, du type *Savignac*, avec *Savigny* ou *Savigné* dans le reste du territoire.

Le suffixe *-ey-*, sous sa forme à vocalisme zéro *-y-*, a une valeur quasi génitivale, un peu d'ailleurs comme la simple voyelle : à partir de *n(e)u* « maintenant » (danois *nu*, angl. *now*, lat. *nu-n-c*) on a un adjectif **new-o* « neuf » ou **now-y-o* avec les deux suffixes de même valeur, dans le nom celtique **Nowio-magos*, littéralement « le nouveau marché » qui a donné *Noyon*. La combinaison *-y-eH_2-*, qui peut se réduire à *-i-H_2* vaut à peu près « celui de... ». Soit le loup, *$w\c{l}k^w o$-* ; *$w\c{l}k^w$-iH_2* sera « celui » ou « celle du loup », peut-être son petit, mais certainement aussi sa femelle. C'est la forme qui a donné le skrt. *vṛkī* et le vieil-isl. *ylgr* « louve ».

La forme réduite *-iH_2* (> *-ī*) du complexe *-y-eH_2* doit être considérée comme la marque traditionnelle du *sexe* féminin. Lorsque, tardivement, apparaîtra un *genre* féminin, marqué initialement dans le pronom démonstratif, puis, par extension, dans l'adjectif correspon-

dant, un suffixe *-eH_2, peut-être déjà sous la forme évoluée *-ā*, va tendre à concurrencer *-iH_2 comme marque de sexe. Mais cette expansion est, en latin, un fait historique : la louve, à l'âge classique, se disait *lupus femina*, la forme en *-a*, l'innovation *lupa*, étant réservée au sens dérivé de « prostituée ». On a déjà signalé que le suffixe de sexe féminin est en latin, /-īk/, comme dans *imperātrīx*, *nutrīx* « nourrice », *uictrīx* « la victorieuse », car le latin, plus souple sur ce point que le français, savait former un féminin sur *uictor*, équivalent de *vainqueur*. Le /-k/ de ces formes a dû apparaître comme un durcissement de -H_2 devant *-s* de nominatif et s'étendre par analogie aux autres cas. La voyelle brève *-i-* est remplacée par la longue sur l'analogie de cas comme l'accusatif où *-iH_2m aboutit à *-*īm*. On a vu que les formes en *-āk-*, issues de *-$eH_2(+s)$, prennent fréquemment le suffixe *-o*, **now-āk-*, par exemple, étant remplacé par **now-āk-o* qui a l'avantage de régulariser la flexion en éliminant un nominatif plus court d'une syllabe que les autres cas : soit, pour « pied », par exemple, lat. *pēs*, en face de *pedem*, *pedis*, etc. De façon parallèle et pour les mêmes raisons, les féminins en *-īk-* vont s'adjoindre le suffixe *-ā* qui apparaît de plus en plus comme l'équivalent féminin de *-o*. C'est à ce type qu'appartient lat. *formīca* « fourmi », féminin, en face de l'équivalent grec *múrmēx* (< *murmāk-s*), masculin, ou lat. *lōrīca* « cuirasse », féminin, en face du grec *thṓrāx*, de même sens, masculin. L'alternance *l-* ~ *th-* rappelle les hésitations *dingua* ou *lingua* pour « langue », *odor* ou *olor* pour « odeur ». De même, sur la racine de *amo* « j'aime », on a *amīca* « amie » qui a dû précéder le masculin *amīcus*. Ce parallélisme existe d'ailleurs dans le cas des formes de base, sans voyelle ajoutée au radical, comme l'atteste, par exemple, *fornāx* « four » en face de *fornix* « voûte » ; à noter, dans ce dernier, la voyelle brève *-i-* dont on n'oubliera pas qu'elle est phonétiquement régulière. On remarque que la forme sans suffixe vocalique se maintient mieux dans les suffixes vivants comme le *-āx* d'*audāx* ou le *-īx* de *nutrīx*.

Le suffixe -iH_2-, avec traitement [k] de la « laryngale », a eu, en slave, une fortune extraordinaire, que ce soit à partir de la forme régulière *-ik-*, ou de la forme à voyelle longue analogique, *-īk-*. Avec palatalisation du *-k-*, il est à la source des innombrables formes en *-ec*, du type russe *otec* « père » [a't'ets], diminutif à partir de la forme enfantine *at(t)a*, ou en [-its], [-itš], des toponymes ou des anthroponymes.

Jusqu'ici la laryngale en cause, lorsqu'il s'agit d'un durcissement en [k], a été constamment notée *H_2, ce qui, bien entendu, peut couvrir plusieurs phonèmes distincts, sourds ou sonores, uvulaires ou pharyngaux. Phonétiquement, c'est naturellement une uvulaire sourde, le *ach*-Laut de l'allemand, qui est la plus proche de l'occlusive dorsale notée [k]. Mais, devant *-s* sourd, une uvulaire sonore s'assourdissait automatiquement.

Dans le mot qui désigne le poisson, comme le vieux-prussien *zuk*, où le vocalisme est toujours *u*, on ne saurait dire s'il s'agit de $*H_2$ ou de $*H_3$. Mais l'élément labial qui s'ajoute à $*H_2$ pour donner $*H_3$ n'a aucune raison d'affecter le durcissement occlusif là où on le postule. On peut donc poser $*\text{-}eH_3 + s > *\text{-}ok\text{-}s$, et *-ōk-s* par analogie, ou peut-être $*\text{-}ak^w\text{-}s$, d'où, ultérieurement, **-ak-s*. On peut expliquer, dans ces conditions, certaines formes du mot qui désigne le corbeau, la corneille ou les espèces voisines : à partir d'une base consonantique $*krH_3$, on a, avec degré *o* de la voyelle entre *k* et *r*, et suffixe *-o*, $*korH_3\text{-}o\text{-}$, où H_3 ne peut affirmer son identité qu'en développant un [w], d'où le latin *coruo-s* « corbeau » ; avec le même degré *o* entre *k* et *r*, la voyelle entre *r* et H_3 et absence de suffixe *o*, on a $*koreH_3(+s)$, d'où le grec *kórax* « corbeau ». Avec degré zéro entre *k* et *r*, la voyelle entre *r* et H_3 et absence de suffixe, on a $*kreH_3(+s)$, d'où, avec une longue analogique, le vieil-anglais *hrōc* « freux », anglais *rook*, le français *freux* lui-même dérivant d'une forme germanique ancienne **χrōk-*. Avec un suffixe en [n], on a le nom de la corneille : grec *korōnē* de $*koreH_3\text{-}n\text{-}\bar{a}$ et, avec d'autres suffixes ajoutés à $*korH_3n\text{-}$, le latin *cornīx* et, finalement *cornicula* qui a donné *corneille*.

Il n'est pas absolument exclu qu'une occlusion glottale, une des deux articulations que nous postulons pour H_1 ait pu voir son point d'articulation transféré de la glotte à la zone dorso-vélaire et, donc, qu'il ait pu y avoir des [k] provenant de $*H_1$ (+s). Mais cette supposition ne paraît guère trouver appui dans les faits. Cf. toutefois le suffixe latin *-fex (-fak-s)* « celui qui fait », par exemple, dans *artifex* « artisan », d'une racine en -H_1, et le *-k-* du verbe *facio* « faire » qui serait analogique de *-fek-s* ou, plus vraisemblablement, du parfait *feci* $< *d^heH\text{-}H_2e\text{-}i$.

Une voyelle ou plusieurs.

A l'issue de ces considérations relatives aux « laryngales », nous ne nous risquerons pas à reconstruire un système de voyelles pour un stade quelconque de l'évolution de l'indo-européen. La réflexion théorique dans laquelle s'insère la théorie des « laryngales » tend évidemment à poser un phonème vocalique unique qu'on pourrait décrire phonétiquement comme un /a/ susceptible de prendre des timbres divers selon les contextes. Les [i] et les [u] des langues attestées ou de nos formes reconstruites ne seraient jamais que des réductions, en syllabes inaccentuées, d'anciennes combinaisons [ay] et [aw]. Les voyelles longues des reconstructions traditionnelles seraient toutes à expliquer comme des allongements de la voyelle en compensation de la chute d'une consonne suivante, « laryngale » ou non, ou comme des coalescences de deux voyelles successives. C'est l'exemple

des langues caucasiques qui a donné aux linguistes le courage d'envisager l'existence de langues à voyelle unique, voire de langues sans phonème vocalique parce que toutes les consonnes y seraient automatiquement accompagnées d'une voyelle d'appui au timbre déterminé par le contexte. Mais un examen attentif et réaliste de ces langues montre que la structure univocalique ne s'impose au chercheur que s'il fait abstraction d'emprunts, de formes marginales, familières, voire argotiques, et s'il n'hésite pas à présenter comme synchroniquement valables les antécédents probables de ce qui est réellement attesté. La conclusion qui s'impose est que, dans tout stade de toute langue, on peut rencontrer des traits qui ne s'intègrent plus pleinement ou qui ne s'intègrent pas encore à ce qu'on peut considérer comme la norme du moment, celle que les linguistes ont bien raison de vouloir cerner, mais sans qu'ils oublient jamais les marges qui l'accompagnent. Pour concrétiser tout ceci, on peut dire qu'il n'y a aucune vraisemblance que les locuteurs de l'indo-européen, à quelque époque qu'on se place, aient jamais eu de difficulté à articuler un [i] et un [u][10]. Si nous reprenons ici la racine qui nous vaut, d'une part, en français *neuf*, en anglais *new*, en allemand *neu*, d'autre part le danois *nu*, l'anglais *now* « maintenant », le latin *nu-n-c*, nous pouvons très bien supposer que, des deux, la forme la plus ancienne est l'adverbe [*nu] avec un [u] bref comme celui du danois *nu*, ou long comme celui du vieil anglais *nū*, et que le **new-* qui est à la source des adjectifs désignant la nouveauté a été refait, à partir de **nu*, sur le modèle de l'alternance *ew* ~ *u* où *u* n'existait que dans les syllabes inaccentuées, alternance encore attestée, par exemple, dans grec *pu-n-th-án-omai* « s'enquérir » en face du futur *peú-s-omai*.

3 ou 4 séries d'occlusives.

Si tout n'avait pas commencé avec le sanskrit, il n'est pas sûr, on l'a déjà suggéré ci-dessus, qu'on aurait posé, pour la langue commune, une série de sourdes aspirées ou, plutôt, la co-existence de deux séries d'aspirées que l'indien est seul à présenter. En fait, les correspondances où apparaissent les sourdes aspirées du sanskrit sont en nombre assez limité. Dans plusieurs cas, la sourde aspirée, *ph*, se trouve dans le groupe initial *sph-*, par exemple dans *sphyá-* « louche de bois » qu'on rapproche de grec *sphḗn* « coin », vieil-anglais *spōn* « copeau », anglais *spoon* « cuillère ». Dans ce cas rien n'empêche de poser au départ une racine en b^h- à laquelle on a antéposé l'*s* mobile qui, sourd lui-même, a fait perdre sa sonorité au b^h- : on aura l'aspirée là où elle existe, en

10. Cf. André MARTINET, Réflexion sur le vocalisme de l'indo-européen commun, *Homenaje a Antonio Tovar*, p. 301-304, repris dans *Evolution...*, p. 108-113.

sanskrit et en grec, la non-aspirée en germanique où, à *bh* sanskrit, correspond *b*, qui deviendra automatiquement *p* après la sourde *s*. Là où *ph-* est initial, dans skrt *phéna-* « écume », par exemple, d'autres langues, comme le vieux-prussien avec *spoayno*, montrent qu'il y a eu alternance de *s-* mobile et de zéro. Dans la grande majorité des cas, les sourdes aspirées indiennes s'expliquent comme des combinaisons de sourdes simples suivies d'un [h] résultant de l'affaiblissement d'une ancienne « laryngale ». On a déjà rencontré ci-dessus, p. 39, *pr̥thu-* « large » rapproché de gr. *platús* sans aspiration ; le cas du mot pour « chemin » a été longuement analysé. On citera ici le mot *roue*, lat. *rota*, de **roteH₂* ; lorsqu'on ajoute le suffixe adjectival *-o-*, on obtient **rotH₂o-* avec chute de la seconde voyelle, d'où, avec une aspirée, skrt. *ratha-* « char », « (engin) à roue ». Le plus souvent le grec ne participe pas à cette aspiration, comme on l'a vu dans *platús*, *póntos*, « mer » et dans la racine *-stā- (hí-stē-mi)*. Mais on en a des traces dans *parth-énos* « vierge » qu'on rapproche de skt. *pr̥th-uka* « jeune animal », et, ce qui retient surtout l'attention, dans la finale skt. *-tha*, grec *-tha* de la deuxième personne du singulier du parfait.

On pourrait donc renoncer à postuler quatre séries différentes pour l'indo-européen commun. Mais alors la question se pose de savoir quels étaient les traits phoniques qui différenciaient, en fait, les trois séries qui finalement s'imposent. Selon la tradition, les séries qui nous restent sont respectivement des sourdes, des sonores simples et des sonores aspirées. Malheureusement, c'est là une combinaison qui n'est attestée nulle part. En effet, on ne connaît de sonores aspirées que dans les langues, comme le sanskrit, qui présentent également des sourdes aspirées. La chose s'explique si l'on pense qu'étant donné la nature de la glotte, combiner les vibrations qui produisent la voix, c'est-à-dire la sonorité, avec le frottement contre ses parois, réclame qu'on dissocie l'action de la partie antérieure de l'organe de celle de la partie postérieure : à l'avant les cordes vocales sont en contact ; à l'arrière elles réservent un orifice pour l'air qui vient des poumons. Cette combinaison délicate peut exister comme combinaison de deux types articulatoires, la sonorité et l'aspiration, utilisées dans le système, mais elle perd sa raison d'être si l'un de ces deux types n'est plus là. S'il n'y a pas, dans la langue, de sourdes aspirées, articulations simples, en dépit de leur dénomination, les sonores aspirées seront, de par la loi de l'économie qu'on appelle le moindre effort, appelées à perdre leur sonorité. Si donc nous éliminons du système les sourdes aspirées, nous devons trouver, pour la série représentée traditionnellement comme $*b^h$, $*d^h$, $*g^h$, $*g^{hw}$, une autre identité.

Les systèmes d'occlusives à trois séries font, le plus souvent, usage de trois positions différentes de la glotte :

1) la glotte à lèvres rapprochées avec des vibrations qui commencent au moment où l'occlusion buccale se relâche ;
2) la glotte solidement fermée dont le relâchement, perçu après celui de l'occlusion buccale, produit ce qu'on appelle la glottalisation ;
3) la glotte ouverte qui, si elle se prolonge légèrement au-delà de l'occlusion buccale, aboutit à l'aspiration.

Dans un système de ce type, la sonorité, c'est-à-dire les vibrations de la glotte, est réservée aux voyelles et à celles des consonnes qui, pour être nettement perçues, réclament le soutien de la voix qui se répercute dans la cavité de résonance, représentée par la bouche, voire les fosses nasales, de forme et de dimension particulières pour chacune d'entre elles. On pense à [n], [m], [l], [r] et a des articulations faibles comme le *v* bilabial (le *v* et le *b* de l'espagnol).

Les trois séries d'occlusives seront donc caractérisées comme 1. à glotte neutre, ou simple, 2. à glotte serrée ou glottalisée, 3. à glotte ouverte ou aspirée. On notera donc, en prenant les apicales en exemple :

1. t
2. tʾ
3. t^{h}

Il peut se faire que, pour la première, les lèvres de la glotte entrent en vibrations avant le relâchement de l'occlusion pour anticiper la voyelle suivante. Il n'y a à cela aucun inconvénient puisqu'il n'y a pas d'occlusives sonores qui doivent rester distinctes de ces simples. Le système pourra, dans ce sens, évoluer vers un état qu'on notera comme :

1. d
2. tʾ
3. t^{h}

Comme les locuteurs ne désirent pas déployer plus d'énergie que nécessaire, on constate, dans un système de ce type, que ce sera tantôt l'aspiration qui s'imposera à l'attention, la glottalisation restant plus discrète, tantôt la glottalisation, l'aspiration n'étant alors qu'esquissée.

Une autre possibilité évolutive sera le relâchement anticipé de la glottalisation qui aboutira à sonoriser une partie de la consonne, d'où passage de [tʾ] à [ʾd] et, s'il n'y a pas d'occlusives sonores, dans le système, l'aboutissement à [d]. On aura donc :

1. t
2. d
3. t^{h}

Lorsqu'on cherche à restituer un sémitique commun, ancêtre des langues qui vont de l'arabe à l'hébreu et à l'akkadien de Mésopotamie, on aboutit à un système occlusif à trois séries où c'est la glottalisation qui a été mise en valeur. Encore attestée en éthiopien, elle a donné l' « emphase » de l'arabe moderne qui, n'étant plus réalisée au niveau de la glotte, mais dans le pharynx, peut se combiner avec la sourdité et la sonorité, ce qui aboutit à un système à quatre séries :

1. d	>	1.	d
2. tʾ	> {	2.	ṭ
		2 bis	ḍ
3. tʰ	>	3.	t⁽ʰ⁾

Les phonèmes que nous avons présentés comme les simples y sont des sonores caractérisées et, ce qui va de pair, l'aspiration des occlusives à glotte ouverte est pratiquement inexistante. Dans les deux cas, il y a anticipation des vibrations de la glotte.

En ce qui concerne l'indo-européen commun, on postule aujourd'hui volontiers, à date très ancienne, l'existence d'une série glottalisée qui, par l'intermédiaire de préglottalisées, a donné les sonores du schéma classique[11]. On partirait de

1. t	pour aboutir à	t
2. tʾ	pour aboutir à	d
3. tʰ	pour aboutir à	tʰ

Ce qui milite en faveur de cette reconstruction est l'inexistence du *b dans les plus anciennes formes reconstruites : en latin, par exemple, *b-* initial n'apparaît que dans des emprunts ou comme le produit d'un ancien *dw- : bonus* provenant de *dwenos, bis* de **dwis,* de la racine de « deux », *duō,* ou encore comme le résultat d'accidents divers : *barba* « barbe », au lieu du **farba* attendu, reproduit à l'initiale le *-b-* de la seconde syllabe ; ailleurs qu'à l'initiale, dans *barba* ou *uerbum* « mot », le *-b-* vient de **dh,* comme l'indique le *d* des équivalents anglais *beard* et *word*. On a signalé précédemment que le germanique ne connaît de *p*, c'est-à-dire l'équivalent d'un **b* primitif, que dans des mots empruntés ou des formations expressives. Or, l'inexistence d'une glottalisée dans l'ordre labial est un trait très fréquent dans les langues qui ont une série de ce type. Dans au moins

11. On trouvera dans Thomas V. GAMKRELIDZE, Language typology and language universals and their implications for the reconstruction of the Indo-European stop-system, *Amsterdam Studies in the Theory and History of Linguistic Science*, IV, *Current Issues in Linguistic Theory*, p. 571-609, un effort très poussé de reconstruction dans le même sens que celle qui est tentée ici, avec une insistance particulière sur la notion de marque.

la moitié des langues en cause citées par Troubetzkoy[12], il n'y a pas de *p*ʾ, et là où on en trouve, ils sont exceptionnels ou statistiquement peu fréquents. La chose s'explique fort bien lorsqu'on examine les conditions de la production des glottalisées. Pour en obtenir une, il faut, en même temps, produire une occlusion dans la bouche et verrouiller la glotte ; ensuite remonter le larynx pour comprimer l'air compris entre les deux fermetures. Cette compression est indispensable pour que le relâchement de la fermeture buccale soit perceptible. Si cette fermeture buccale est profonde, par exemple au niveau uvulaire, le volume d'air compris entre les deux occlusions est faible, si bien que la montée de la glotte aboutit vite à réaliser la pression nécessaire. Si la fermeture buccale est au niveau du palais dur ou des gencives supérieures, la cavité est plus vaste et la pression obtenue pour la même montée du larynx est moindre. Mais les parois de cette cavité restent fermes puisque s'appuyant sur des os ou représentées par la langue qui est un organe musclé. Mais si, maintenant, la fermeture buccale est réalisée par les lèvres, le volume compris entre ces organes et la glotte est considérable et, circonstance aggravante, les joues, de faible musculature, font partie des parois de la cavité en cause. Réaliser une glottalisée dans ces conditions devient presque une gageure.

Comme le passage des glottalisées aux sonores, par l'intermédiaire des préglottalisées, est un phénomène bien identifié, il est donc tentant de poser, à un stade ancien de l'indo-européen commun, un passage, dans notre série 2, de *t*ʾ à *d*. Si nous présentons le système occlusif entier, on aura donc passage de :

1.	p	t	k	k^w	à	1.	p	t	k	k^w
2.		tʾ	kʾ	$k^{ʔw}$		2.		d	g	g^w
3.	p^h	t^h	k^h	k^{hw}		3.	p^h	t^h	k^h	k^{hw}

Si nous n'avions à rendre compte que du consonantisme du latin et du grec, nous pourrions allègrement partir de ce dernier schéma. En latin, par exemple, la série 2 serait attestée telle quelle à l'initiale de *duō* « deux », *genus* « genre », à l'intérieur d'*inguen* « aine », avec affaiblissement de *gu* en *u* à l'initiale dans *uenio* « venir » (cf. got *giman* = k^w*iman*) et création, nous l'avons vu, d'un *b* à partir de [dw-]. La série 1. serait représentée dans *pater* « père », *trēs* « trois », *cor* « cœur », *quod* « que ». Dans la série 3. l'occlusion se relâche, ce qui donne, dans un premier temps, *f*, *Þ*, χ et χ^w, la série se réduisant finalement à *f* et *h*, par élimination de *Þ* et χ^w au profit de *f*, ceci ne valant d'ailleurs qu'à l'initiale, les produits internes étant diversement affectés par les contextes. On a donc à l'initiale *fero* « je porte », grec *phérō* (angl. *bear*), *facio* « je fais », gr. *tí-thē-mi* « je place » (angl. *do*),

12. Dans *Principes de phonologie*, Paris, 1949, p. 165 et suiv.

hiems « hiver », gr. *kheimôn* « hiver », *khiỗn* « neige », *formus* « chaud », gr. *thermós* (avec *th* régulier pour une aspirée labiodentale devant *e* ; angl. *warm*).

Mais les formes anglaises, qu'on vient d'ajouter aux équivalents grecs des mots à aspirées initiales, suggèrent bien que la chose n'est pas si simple. Une série 3. p^h, t^h, k^h, k^{hw} n'explique pas les *b*, *d*, *g*, *w* du germanique, les *b*, *d*, *g* du celtique, les *b*, *d*, *z*, *g* du slave, tous uniformément sonores. Nous sommes donc ramenés à poser, pour la série 3., *à un certain stade,* des phonèmes combinant aspiration et sonorité, c'est-à-dire ceux du sanskrit et de la tradition comparatiste. Mais ceci nous entraîne à poser, *pour ce stade*, une quatrième série de sourdes aspirées, nées de la combinaison d'un phonème de notre série 3 avec un [h] suivant, résidu d'une « laryngale » sourde plus ancienne. Un complexe [t + h], par exemple, ne se serait pas confondu avec t^h, mais, représentant une articulation glottale plus énergique, aurait tendu à affaiblir le t^h en une sonore faible pour mieux contraster les deux types. Seul l'indo-iranien aurait préservé les deux séries d'aspirées fortes ($t + h > {}^*t^h$) et faibles (${}^*t^h > {}^*d^h$). L'italique, latin compris, et le grec auraient confondu les deux types sous la forme t^h. Les langues de l'Europe du nord, de leur côté, auraient assimilé ${}^*t^h$ à leur **t* et finalement réduit ${}^*d^h$ à [d]. Ce dernier ne s'y est pas nécessairement confondu avec le produit du **d* ancien (< **tʼ*), comme on le voit par le germanique qui a *d* pour ${}^*d^h$ (angl. *do*) mais *t* pour **d* (angl. *two* « deux »). Noter également que le celtique à *b* pour ${}^*g^w$ (de notre série 2) dans irl. *bó* « vache », mais *g* pour ${}^*g^{hw}$ (de notre série 3), dans irl. *gonim* « je frappe ».

Comme le hittite conserve au moins certaines des anciennes laryngales sous la forme de consonnes bien caractérisées, et non comme simple aspiration en dépit des translitérations en *h*, on peut se demander s'il a, dans sa préhistoire, participé au processus de création d'une nouvelle série d'aspirées par combinaison d'une simple et d'un [h]. En tout état de cause, le hittite, s'il semble distinguer entre notre série 1 et les deux autres, n'a pas trace d'une différenciation entre 2 et 3. L'opposition entre 1 d'une part, 2 et 3 réunis, d'autre part, serait, selon l'interprétation courante des graphies, celle de fortes à des faibles.

On peut résumer comme suit le développement qu'on vient de postuler pour le système occlusif de l'indo-européen :

	I		II		III		IV
1.	t	>	t	>	t	>	t
2.	tʼ		d		d		d
3.	t^h		t^h		t^h		d^h
4.					t + h		t^h

Le système IV est celui du sanskrit. Pour le grec et l'italique, on posera, à partir de IV, une confusion des séries 3 et 4 ; pour les autres langues européennes, l'élimination probable de 4 au profit de 1.

Les occlusives complexes.

Il y a tout un ordre d'occlusives qu'il a fallu envisager dans l'opération de reconstruction et dont nous n'avons pas fait état jusqu'ici. Il s'agit d'articulations complexes dont on ne sait trop comment reconstruire la forme primitive. On la postule à partir de correspondances comme celles du nom de l'ours. Le sanskrit a pour désigner ce plantigrade la forme *ṛkša-*, le grec *árkto-*, le latin *ursu-s* (d'**urcso-s*), le gallois, *arth* (d'**ark'o-*) représentant le celtique. La forme celtique nous amène à poser une forme débutant en *H_2er-* alternant avec *H_2ṛ-* qui est postulé par le sanskrit et le latin. L'alternance de *t* en grec et en celtique et d'une sifflante ou chuintante en sanskrit et en latin n'est attestée ailleurs que dans un complexe commençant par une occlusion vélaire. Elle apparaît dans des mots aussi bien établis que les désignations de la hache et de l'œil. Des complexes analogues sont attestés dans notre série 3, avec la désignation de la terre, en sanskrit *kṣam-*, en grec *khthōn*. Dans les reconstructions traditionnelles, on pose, comme second élément du complexe, une interdentale sourde ou sonore qui est le choix de gens qui constataient que les étrangers, mis en face du *th* anglais de *thin*, hésitaient entre *sin* et *tin*. Mais les comparatistes avisés n'ont jamais vu, dans le *þ* utilisé dans ce cas, qu'une convention graphique correspondant à une réalité phonique indéterminée.

Il y a, dans l'indo-européen reconstruit, d'autres phonèmes correspondant à des articulations complexes. Quand nous posons un **k^w*, nous supposons que les organes réalisent *en même temps* l'articulation du [k] et celle du [w], mais l'élément [w] tend à s'imposer à l'attention non seulement au moment précis du relâchement de l'occlusion, mais un peu après, d'où l'évolution du latin *quattuor*, où l'on suppose encore un [k^w], à l'espagnol *cuatro* où l'on a une succession [k] + [w]. Dans bien des langues, en Afrique occidentale notamment, on trouve des phonèmes complexes qui combinent l'articulation de [k] et celle de [p] avec relâchement successif de la fermeture de [k] et celle de [p], ce qui se comprend bien, puisque l'air qui monte des poumons va rencontrer d'abord la fermeture dorso-vélaire, ensuite la fermeture labiale. On trouve donc un ordre de labio-vélaires notées /kp/, /gb/. On voit très bien comment ces dernières peuvent se relâcher en k^w et g^w, si bien que rien n'empêche de penser que le /k^w/ du latin *quis* remonte en fait à un [kpis]. Le passage de **$k^w$is* à *pis* qu'on constate en osque, par exemple, n'en serait que plus facile, [kpi] passant à [pi] par simple

relâchement du [k]. L'exemple du roumain *apa* « eau » qui vient du latin *aqua* suggère sans doute qu'on peut passer directement de [k^w] à [p], mais qui nous dit que l'intermédiaire n'est pas précisément [kp] ?

Si, après cet excursus, nous revenons au cas des mots pour « ours » et « terre », on pourrait se demander s'il ne faut pas poser, au départ des correspondances *khth, kt/kš*, des phonèmes complexes de type [kt], [gd] distincts des groupes **kt* et **gd* qui, eux, résulteraient de la rencontre de deux occlusives après chute de voyelle inaccentuée intermédiaire. Cela voudrait dire que dans une forme comme **okto* « huit », une trace vocalique est longtemps demeurée entre **k* et **-t-*, donnant quelque chose comme *ok*ᵊ*tō*.

Il faut noter que ces phonèmes complexes, pour lesquels on vient de suggérer des formes de type [kt], apparaissent, ailleurs qu'en sanskrit et en grec, sous des formes simplifiées. Nous avons déjà vu le *-s-* du latin *ursus* et le *th* du gallois *arth*, mais la chose est aussi nette du côté de notre série 3 d'aspirées. Le phonème pour lequel nous pourrions poser g^{dh} qui se trouve à l'initiale de la désignation de la terre ne conserve sa forme complexe que dans le skt *kṣam-* et le grec *khthṓn ;* partout ailleurs on trouve le reflet soit d'un *g^h, soit d'un *d^h. En grec même, « à terre » se dit *khamaí* et en lat. *humī*, formes au départ identiques. Pour « terre », le latin a *humus ;* avec *z* ou *ž* provenant régulièrement de *g^h dans les langues *satem*, on a le lit. *žẽmė*, le vieux-prussien *semme* (= *zeme*), le vieux-slave *zemlja ;* l'homme, qui est l'être terrestre par opposition aux dieux, apparaît, avec un suffixe *-en-*, en latin comme *homō*, gén. *hominis*, en vieil-angl. comme *guma* (déformé en *-groom*, dans *bridegroom* « le marié »). Correspondant au grec *khthṓn*, l'irlandais à *dū* « endroit », accusatif *don*, avec *duine* pour « homme ».

La même alternance de *g^{dh} et *g^h apparaît dans la désignation du poisson (cf. ci-dessus) avec la forme complexe dans gr. *i-khthû-s* et la forme simplifiée dans les formes baltes *žuvis, zuk-, žuk-* et arménienne *jukn*.

Les sifflantes.

Le traitement *kš-* du sanskrit — qui, on a pu le constater, vaut aussi bien pour *g^{dh} que pour *k^t — rappelle le comportement inattendu d'un groupe **tt* résultant du rapprochement d'une apicale finale de la racine et du *t* initial d'un suffixe. Ce groupe ne se maintient guère comme tel qu'en sanskrit où on le considère comme dû à des restaurations analogiques. Partout ailleurs, il se trouve représenté soit par *-st-*, soit par *-ss-*. De la racine de « voir » avec la valeur de « savoir », *weid-/wid*, et du suffixe *-tōr*, on a en grec *(w)ístōr* « celui qui sait ». En latin, le participe passé de *pat-ior* « je souffre » est *passus*

(< *pat* + *tos*). Mais, là encore, il a dû y avoir des restaurations analogiques soit de la première consonne du suffixe, d'où *-st-*, soit de la dernière consonne de la racine, d'où *-ts-* évoluant en *-ss-*. En fait, on a tout lieu de penser qu'à un certain stade de l'évolution, la géminée était remplacée par la simple : le *-*ss*- attendu à la 2e personne du singulier du verbe « être », *es* + la désinence *-si*, d'où **essi*, est attesté en grec comme *ei*, c'est-à-dire un plus ancien **esi*, avec chute régulière de *-s-* entre voyelles. On voit mal ce qui a pu déterminer ce passage de *-*tt*- à *-s-*, sans doute par l'intermédiaire de *-*ss*-. Mais il ne fait aucun doute qu'il y a eu passage d'une articulation occlusive à une sifflante, même si les causes du phénomène nous échappent. On verra plus loin les traits morphologiques que cette évolution permet d'expliquer.

Dans le même ordre d'idées, il n'est pas exclu de penser que le *s* unique des reconstructions traditionnelles recouvre peut-être différents types phonologiques anciens. La plupart des langues du monde où l'on rencontre les trois séries d'occlusives que nous avons présentées comme **t*, $*t^{ʔ}$ et $*t^{h}$, ont à côté des ordres des types [p], [t], [k] strictement occlusifs, des ordres affriqués divers et notamment, pour ne citer que les sourdes, des *ts*, des *tš*, voire des *tl-*. On pourrait donc penser que le **s* unique de nos reconstructions remonte à un ordre de trois phonèmes que, si nous notons l'affriquée [ts] au moyen de *c*, nous reconstruisons comme **c*, $*c^{ʔ}$ et c^{h}. Si la réduction de ces affriquées à la simple fricative *s* date d'une époque où les glottalisées n'étaient pas encore passées à des sonores, on comprend qu'il y ait eu confusion des trois unités en un phonème unique, puisque les articulations glottalisées et aspirées caractérisent des articulations occlusives ou à noyau occlusif, mais non des fricatives qui, à moins qu'elles soient sonores, réclament une glotte grande ouverte.

Cette restitution d'un ordre sifflant expliquerait la fréquence de *s* aussi bien dans les diverses langues que dans les reconstructions. Dans la simple flexion nominale on verrait assez bien qu'à date ancienne, on ait pu distinguer entre la particule qui allait fournir la marque du nominatif et du génitif d'une part, et d'autre part l'élément qui va servir à marquer le pluriel. Certains traits du hittite pourraient peut-être venir à l'appui d'une telle hypothèse, qui, il faut le reconnaître, se fonde essentiellement sur l'expérience que nous avons aujourd'hui des systèmes phonologiques en général. Dans toutes les langues traditionnellement utilisées par les comparatistes, tous les *s*, dans un contexte donné, ont le même comportement.

Les sonantes.

On n'a, jusqu'ici, traité des liquides et des nasales que pour signaler qu'elles appartenaient à la catégorie des sonantes, c'est-à-dire de ces

phonèmes qui, à l'instar de *y* et de *w*, fonctionnaient soit comme consonnes, soit comme « voyelles », ou plus exactement comme centres de syllabes : lorsque la syllabe était accentuée, elle présentait la voyelle, celle qui apparaît plus tard sous la forme de *e* ou de *o ;* si elle perdait son accent, la voyelle disparaissait et, si elle était suivie, dans la syllabe, d'une nasale, d'une liquide *l* ou *r*, d'un *y* ou d'un *w*, c'est ce dernier qui assumait désormais la fonction de centre de syllabe, noté respectivement *m̥*, *n̥*, *l̥*, *r̥*, *i* et *u*. Ce qui pouvait donc se voir affecté était leur identité comme consonnes. Mais ces sonantes étaient postulées bien distinctes les unes des autres.

En ce qui concerne la distinction entre *l* et *r*, on n'a pas d'indication qu'elle ait jamais, à date ancienne, été menacée. Il y a, de par le monde, beaucoup de langues où la latérale et la vibrante se distinguent mal ou pas du tout : un Japonais reproduira *villa* comme *bira*. Mais, parmi les vieilles langues indo-européennes, seul le sanskrit manifeste des traces d'instabilité de cette opposition.

En ce qui touche aux nasales, on ne met guère en doute leur identité respective. Il est certes fréquent que ce qu'on reconstruit comme **-m* final soit attesté comme *-n*. C'est le cas, par exemple, de la finale de l'accusatif singulier et du génitif pluriel en grec. Mais on n'a pas de raison de postuler, pour la langue commune, une hésitation entre la labiale et l'apicale. A l'intérieur du mot, on constate que la qualité, *m* ou *n*, voire [ŋ], de la nasale est souvent sous la dépendance du contexte consonantique. Il existe, dans la conjugaison, ce qu'on appelle l'infixe nasal, élément qui peut apparaître avant la dernière consonne de la base, par exemple, en grec, dans le présent *la-m-bánō* « je prends » en face de *é-lab-on* « je pris », en latin, dans *li-n-quo* « je laisse » en face de *liquī* « j'ai laissé ». Mais on se rappelle peut-être le lituanien *šim̃tas* « cent » qui, même devant *-t-* qui laisse attendre un *-n-*, comme dans le latin *centum*, conserve fidèlement le *-m* final de *dekm̥* « dix », devant le préfixe *-to-* dans le dérivé *(d)km̥-to-m*.

Les prénasalisées.

Ce qu'il faut relever, c'est que l'on rencontre très fréquemment un complexe *-nt-* dont on ne constate pas qu'il soit jamais le produit de la chute d'une voyelle ancienne entre *-n-* et *-t-*, ou, comme dans *centum*, le résultat d'un contact entre finale de base et initiale de suffixe. En d'autres termes, *-nt-* paraît se comporter plutôt comme une unité que comme un groupe. Sa fréquence d'ailleurs paraît également avoir été celle d'un phonème et non celle d'un groupe de phonèmes. Or, il est fréquent, un peu partout dans le monde, que, dans chaque ordre de consonnes, figure à côté d'une non-nasale, disons, par exemple, /d/, et d'une nasale /n/, une « prénasalisée » /ⁿd/. On connaît le fréquent

patronyme sénégalais noté *N'Diaye,* le nom de la capitale du Tchad *Ndjamena* avec son [^{n}dž] initial prénasalisé, celui du directeur de l'Unesco, *M'Bow.* Les graphies avec apostrophe entre la nasale et l'occlusive suivante suggèrent la chute d'une voyelle, ce qui, naturellement, n'est pas le cas.

Mais ce phénomène n'est nullement limité à l'Afrique. Le témoignage du basque *makila* « bâton », évidemment dérivé du latin *bacillum,* plur. *bacilla,* amène à poser, pour le basque ancien, un /^{m}b/ qui a paru plus adéquat pour reproduire le *b-* latin que le *p-* faible ou le *p-* aspiré qu'on peut restituer pour l'euskarien par le rapprochement des données diverses. Il faut certainement envisager un /^{m}b/, assez généralement, dans le bassin méditerranéen, d'où, par exemple, l'hésitation entre *sambūcus* et *sabūcus* pour la désignation latine du sureau.

On pourrait donc envisager l'existence, à un stade ancien de l'indo-européen, de prénasalisées, au moins dans l'ordre apical, surtout pour la série n° 1, donc /^{n}t/ et aussi pour la série numéro 2, d'où /^{n}t ʔ/ passé à /^{n}d/. Ces articulations s'ajouteraient aux autres phonèmes complexes, les labiovélarisées traditionnelles, et ceux qui sont attestés dans les désignations de l'ours et de la terre.

Les prénasalisées existent aussi bien à l'initiale qu'à l'intervocalique et c'est cette latitude qui fait soupçonner leur caractère de phonème unique : aucun Européen ne perçoit le nom *Kilimandjaro* du point culminant de l'Afrique comme *ki-li-ma-ndja-ro,* ce qui est pourtant, selon toute vraisemblance, la coupe syllabique correcte. On voit mal ce qui serait, dans les langues indo-européennes attestées, le produit initial d'un *^{n}*t-* ancien : *n-* initial n'est pas particulièrement fréquent dans le lexique des langues de la famille, même s'il n'est pas rare dans les textes du fait de la négation. Une évolution *^{n}*t* > [t-] n'est pas impensable, mais on cherche en vain les données qui en justifieraient la reconstruction.

L'alternance -r/-n-.

L'intérêt du rappel du statut particulier de *-nt-* en indo-européen se trouve dans la possibilité qu'il suggère d'expliquer la curieuse alternance de *-r* final et de *-n-* dans un certain nombre de noms neutres très largement attestée, tout particulièrement en hittite où l'on soupçonne qu'il s'agissait, non pas, comme ailleurs, d'un résidu en voie d'élimination, mais d'un procédé vivant de dérivation. L'illustration la plus claire nous est probablement fournie par le latin *femur* « cuisse » qui a cette forme au nominatif-accusatif indifférencié ; aux autres cas, on trouve un radical *femin-,* au génitif *feminis,* par exemple. Mais l'analogie a joué dans les deux sens et l'on trouve un nominatif-

accusatif *femen* et inversement un génitif *femoris* qui, finalement, s'établit comme la norme ; d'où l'adjectif *femorālis* « de la cuisse ».

On posera qu'à un stade ancien, *-*n* final est passé à -*r* sous la pression de *-^{n}*t* tendant vers -*n*. Entre deux voyelles *-^{n}*t*- susceptible de passer à -*nt*- ne posait pas de problème. Pour les noms, seuls les neutres pouvaient présenter l'*n* ancien à la finale. Les animés, masculins et féminins à venir, présentaient toujours une désinence, et notamment -*m* à l'accusatif, -*s* au nominatif. L'-*n*- de fin de radical n'apparaissait à la finale du mot que dans des stades récents où, le groupe final *n-s* du nominatif ayant finalement été éliminé par une simplification du groupe, suivie d'une nasalisation éphémère de la voyelle précédente, la consonne nasale avait été restituée sur l'analogie des autres cas.

Il y a, certes, beaucoup de neutres attestés qui se terminent en -*n*. En latin, par exemple, où l'on trouve *flumen* « fleuve », *carmen* « chant » et bien d'autres. Ce -*men* remonte à *-*ment*, plus anciennement *-*mn̥t*, avec *-^{n}*t* réduit à -*n* au moment où *-*n* passait à -*r*. On rapprochera de *carmen* le nom *Carmenta* (couper *Carment-a*) d'une nymphe romaine.

Dans ces neutres latins, à nominatif-accusatif très fréquent, l'analogie a imposé un -*n*- aux autres cas, donc au génitif, *fluminis, carminis*. Mais les équivalents grecs, en -*ma* (pour *-*mn̥*), ont au génitif -*matos* (de *-*mn̥t*-) où le *-^{n}*t*- original a été préservé par la voyelle suivante. En latin même, lorsque le suffixe -*o* s'est rajouté aux neutres, le -*t*- n'avait pas disparu, d'où la forme -*mentu*- qui a finalement remplacé l'ancien -*men* et qui, sous la forme -*ment*, reste un des éléments les plus productifs de la dérivation française. Dans les non-neutres, les nominatifs en -*ont*- ont évolué de façon variée : on a -*ōn* en grec dans les participes présents, -*ons* en latin, prononcé [-ōs] dans *pons* « pont » par exemple.

Un autre mot latin en -*er*, *iter* « chemin », nous met sur la voie d'un autre produit de l'évolution de *-*n* final à -*r*. Ce terme a, au génitif, *itineris*, qui alterne avec *iteris*, ce qui permet de poser un ancien **itinis*. Or, *iter* est, avec ancien accent sur la finale, **itér*, un doublet de *éitr̥* qui, a donné le latin *ītur* « on va », « il y a du chemin », ce qui suggère que le -*r*, qui caractérise les impersonnels, puis, par extension, les passifs en italique et en celtique, remonte en dernière analyse à un -*n* final.

On résumera ce qui précède par le tableau suivant :

*-*n* → -*r*	*-^{n}*t* → -*n*
*-*n-o*- → -*no*-	*-^{n}*to* → -*nto*
*-*on* + *s* → -*ō* → *ō*	*-*o*^{n}*t-s* → -*onts* → -*ō*
	-*ons*
	-*ōs*

L'alternance -$*b^h$-/-m-.

Dans les langues où l'on trouve des occlusives prénasalisées, elles sont normalement représentées dans les différents ordres, labial, apical, dorso-palatal, dorso-vélaire. Elles forment donc une série du type $^m b, ^n d, ^n g$. Il n'est pas rare qu'elles se présentent sous la forme de deux séries, par exemple, une série sonore, comme la précédente, et une série sourde $^m p$, $^n t$, etc. On peut donc envisager, en indo-européen, à côté de $*^n t$, un $^m p$ d'une part, un $*^n d^h$ d'autre part [13], donc un système de prénasalisées du type :

$*^m p$	$*^n t$	...	...
..	$*^n d^h$	...	...

A partir d'un $*^m b^h$ ancien, on pourrait expliquer les formes variables qu'assument certaines désinences du pluriel des noms : au datif, à l'ablatif, à l'instrumental, on y rencontre, dans la plupart des langues anciennes, des formes qu'on reconstruit avec $*$-b^h-, alors qu'aux mêmes cas, on trouve un *-m-*, en germanique, en balte et en slave. Qu'on compare le latin *nō-bis* « à nous » et l'équivalent russe *na-m*. On pourrait donc poser, au départ, des formes en $*^m b^h$ où la nasalité se serait imposée dans les langues du nord et aurait disparu ailleurs. Ces diverses désinences, en sanskrit, par exemple, le datif-ablatif *-bhyas* et l'instrumental *-bhis*, en russe, respectivement *-m* et *-mi*, doivent représenter une ancienne particule qu'on reconstruit sous la forme $*b^h i$ et qui est attestée en grec sous la forme *phi*, par exemple, dans *autophi* comme cas oblique de *autós* « soi-même ». Lorsqu'à $*b^h i$ s'ajoute une voyelle *e/o*, le *i* peut se maintenir comme *-y-*, mais il peut aussi disparaître comme dans la finale *-bus* (< $*$*-bos*) du latin (cf. *omni-bus*). En germanique, elle est attestée sous la forme du préfixe *be-*, all. *besprechen*, angl. *bespeak*, qui a notamment pour fonction de transitiviser les intransitifs : *sprechen* « parler », *besprechen* « traiter de, commenter », mais également comme préposition et comme préfixe : all. *bei*, angl. *by*, qui marque la proximité, l'adjonction, all. *beisteuern* « contribuer » (*Steuer* « taxe, impôt »). Si notre hypothèse est admise, on reconstruira donc $*^m b^h i$. Avec un élément préfixé et la valeur d' « autour », on peut poser $*H_2 e$-$^m b^h i$ qui donne en grec *amphi* et, en germanique, avec vocalisme zéro du premier élément, $*H_2{}^m b^h i > *H_2 \c{m} b^h i$, d'où vieil-anglais *ymb*, all. *um*. *Amphi* a également la valeur de « des deux côtés de » et doit être rapproché du latin *ambō* « l'un et l'autre », avec le suffixe *-ō* de duel; pour cette notion, le germanique présente la même base, sans élément H_2e- préfixé et le

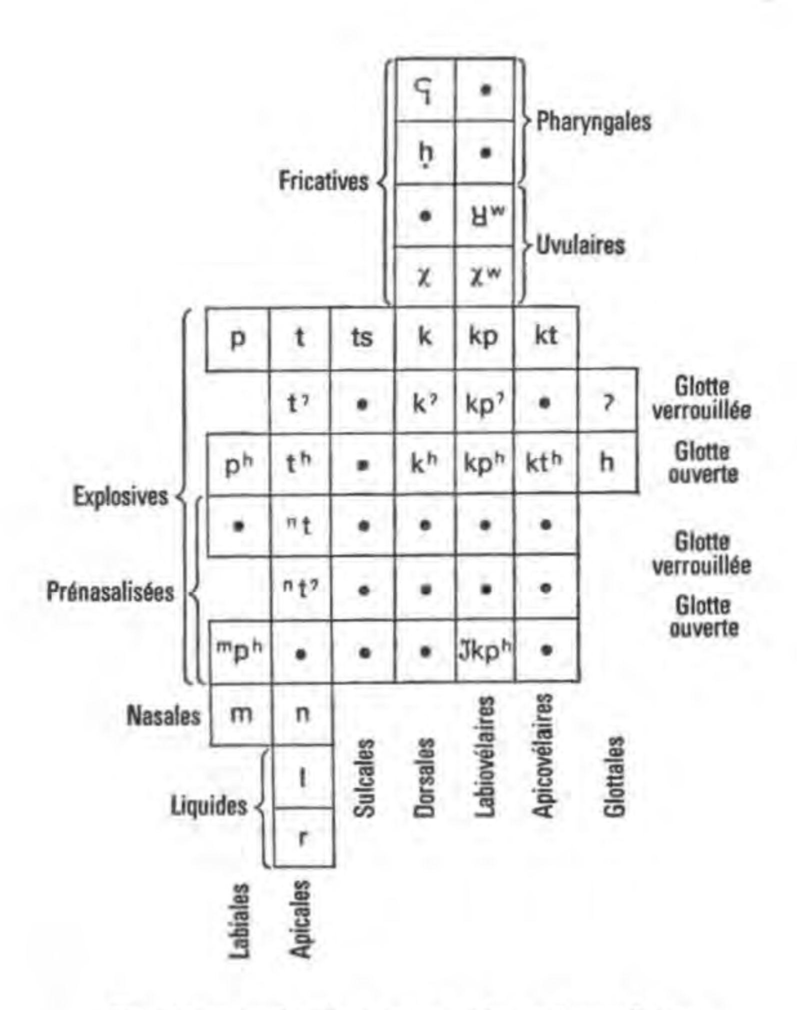
13. Voire un $^n g^{hw}$ dans le prototype pour *éponge*; cf., ci-dessus, chapitre V, note 2, p. 96

traitement $*b^h$ du $*^m b^h$- initial, dans all. *beide*, angl. *both*, de $*b^hei$ + une forme du démonstratif, dan. *begge*, de $*b^hei$-*yo*-.

On ne doit pas s'étonner de trouver, selon les langues et les positions, initiale, intervocalique et finale, des traitements différents des prénasalisées que nous postulons ici. De façon générale, la prénasalisation se perdrait à l'initiale ($*^m b^h i > b^h i$) et s'imposerait à la finale aux dépens de l'occlusion orale (*-*nt* > -*n*), l'élément le moins central étant éliminé dans les deux cas. A l'intervocalique il y aurait maintien des deux éléments et, finalement, traitement du complexe comme deux phonèmes successifs avec une coupe syllabique du type **ten-te* au lieu du $*te\text{-}^n te$ plus ancien. Il est vraisemblable que certains des -*nt*- attestés remontent à d'anciennes successions -*net*- avec chute du -*e*-, ce qui veut dire que les **-ten-te-* issus de **te-ne-te* se seraient confondus avec ceux qui proviendraient de $*te\text{-}^n te$. La chose ne saurait étonner. C'est un fait bien établi que, dans les langues indo-européennes attestées, on trouve confondus d'anciens phonèmes uniques et le résultat de combinaisons de phonèmes successifs. En latin, par exemple, *qu* correspond aussi bien à un ancien $*k^w$ qu'à une combinaison de $*k + u$ devant voyelle suivante : -*que* « et » correspond au grec -*te* et à un germanique ancien -*h*, c'est-à-dire à $*k^w e$, tandis que *equos* « cheval » se révèle comme dérivé d'un ancien **eku-* « rapide » (cf., peut-être, le grec *ōkús* même sens) suivi du suffixe adjectival -*o* ; de même, pour lat. *inguen* « aine », rapproché du grec *adēn*, de même sens, on reconstruit $*\d{n}g^w en$- avec $*g^w$, mais dans *lingua* « langue », on a une ancienne succession $-g^h + u + eH_2$, comme nous l'avons vu ci-dessus.

La différence de traitement du $*^m b^h$ dans les désinences issues de $*^m b^h i$, $*b^h$ ici et *m* là, pourrait s'expliquer par le maintien, dans les langues méridionales, d'une certaine autonomie de $*^m b^h i$ que rappelle le comportement du produit grec *phi*, d'où le traitement $*b^h i$, attendu à l'initiale. Dans les langues septentrionales qui présentent -*m*, on peut penser à une évolution plus rapide de $*^m b^h i$ vers le statut de désinence avec chute du -*i*, celle qu'il faut bien envisager pour expliquer son absence dans le datif russe *na-m*, ainsi d'ailleurs que dans la désinence latine -*bus*, de $*\text{-}b^h os$. On peut expliciter les divers traitements comme suit :

$*^m b^h i \rightarrow *b^h i$	comme adverbe > préposition : all. *bei*, angl. *by*.
$\rightarrow *b^h i$	comme particule non amalgamée : grec *phi*.
$\rightarrow *\text{-}b^h i$	comme désinence tardivement amalgamée : skrt. -*bhy-as*, -*bhi-s*.
$\rightarrow$ -*m*	comme forme anciennement amalgamée, coexistant avec $*b^h i$ comme particule non amalgamée, avec extension analogique de -*m*, dans les langues du nord, de $*b^h$, dans une langue comme le latin.



Phonèmes envisagés pour une très ancienne forme
de l'indo-européen

Un point dans une case indique un phonème possible qui n'a pas été explicitement postulé

Ce qui, dans ce tableau, réclame une explication est le rangement, dans les mêmes colonnes, d'occlusives et de « laryngales ». Les phonèmes placés dans la colonne « dorsales » sont effectivement articulés avec le « dos » de la langue, c'est-à-dire sa face supérieure, mais, pour les occlusives, contre le voile du palais, pour les fricatives, dans la direction de la luette et de la paroi postérieure du pharynx, donc dans des parties différentes du chenal. On a fait abstraction de ces variations en ne retenant comme distinctive que l'opposition d'occlusive à fricative. Tout cela vaut également pour les labio-vélaires. Le choix, dans ce cas, de [p] ou de [ʷ] comme indicateur du trait de labiovélarité va de pair avec occlusion et friction respectivement. Pour les pharyngales et les uvulaires, la première tranche horizontale correspond à des sonores, la deuxième à des sourdes.

L'accent.

Avant d'en finir avec la phonologie, il nous faut revenir sur le problème de l'accent. Nous avons, dans ce qui précède, opéré constamment avec la notion d'un accent qui, à un certain stade de l'évolution de la langue, a eu pour effet de préserver la voyelle accentuée et de faire disparaître ou, au moins, d'affaiblir celles qui ne l'étaient pas. Là-dessus s'est greffé le maintien de ce qu'il fallait de voyelles pour rendre les consonnes prononçables si elles ne pouvaient pas, elles-mêmes, assumer la fonction de centre de syllabe. Puis, ultérieurement, lorsque les conditions accentuelles étaient différentes, il y a eu des rétablissements analogiques de voyelles.

Ce type d'accent reconstruit, très ancien, a été un peu partout remplacé par autre chose. On ne peut faire totalement confiance ni au sanskrit, ni au grec, ni au latin, pour nous renseigner sur la place et la nature de l'accent qui a abouti à l'hécatombe de voyelles qu'il faut postuler. C'est peut-être le germanique qui est resté le plus longtemps fidèle au type primitif, comme l'indiquent les alternances couvertes par la Loi de Verner, cf., ci-dessus, p. 88 : l'alternance allemande de *-h* et de *-g-*, dans *ziehen* « tirer » et *ge-zogen* « tiré », nous indique que l'accent était, au présent, sur la racine dans **deuko-* (comparer le latin *dūcō* « conduire » avec un *ū* long issu d'une diphtongue) et au participe sur le suffixe dans **duk-ón-* (lat. *ductu-s* avec un autre suffixe) et ceci est confirmé par *-eu-* dans un cas, *-u-* dans l'autre.

Cet accent, lorsqu'il a eu les effets que l'on constate, avait-il sa place déterminée par un principe général, ce qui est souvent le cas : première syllabe du mot en tchèque et en hongrois, avant-dernière syllabe en polonais, place de l'accent réglée par la quantité de l'avant-dernière syllabe en latin, dans tous les cas où l'on parle d'un accent fixe ? Ou cette place variait-elle de forme à forme ? On serait tenté, sur la foi de certaines données, de diagnostiquer, pour une certaine époque, une tendance à placer l'accent sur le dernier élément du mot, racine ou suffixe, avec certaines désinences attirant l'accent et d'autres le laissant sur ce qui précède. Nous retrouverons tout cela ci-dessous lorsque nous traiterons de la grammaire.

On se contentera de rappeler ici que la conception d'un accent, qui, comme un *deus ex machina,* descend du ciel pour mettre à mal le vocalisme d'une langue, est une vision romantique qui n'a rien à faire avec la réalité. Si on a eu l'air de s'en inspirer ci-dessus, c'est simplement parce que, dans sa naïveté, elle permet un exposé plus simple des faits. En réalité, il y a des éléments du discours dont la valeur informatrice est faible et, dans la mesure où ils se concentrent dans une partie du discours caractérisable en termes formels, ils vont

tendre à s'affaiblir, l'énergie nécessaire à leur production se transférant en grande partie sur leurs voisins plus favorisés. En d'autres termes, ce n'est pas l'accent fort qui élimine les voyelles, mais l'affaiblissement des voyelles qui contribue à renforcer celles qui demeurent. La fonction propre de l'accent, c'est l'établissement de contrastes entre les différentes parties successives de l'énoncé, pour mieux les distinguer les unes les autres et mettre certaines d'entre elles en valeur. La fonction primaire de l'accent est démarcative ou, au moins, culminative en ce qu'il crée dans l'énoncé des sommets qui signalent la présence d'un élément d'information central.

Les tons.

L'accent n'a en principe rien à voir avec les tons qui sont un comportement particulier de la courbe mélodique du discours sur le noyau vocalique central d'une syllabe. Ce comportement permet de distinguer, dans les langues à tons, deux syllabes composées des mêmes phonèmes, vocaliques et consonantiques, disons, en chinois, *mā* « mère », de *má* « chanvre », *mǎ* « cheval », *mà* « injurier ». Il se trouve seulement que si, dans une langue à tons, se développe un accent, ce pourra être aux dépens de la netteté des timbres des voyelles inaccentuées et aussi de celle de leurs tons. Il n'est donc pas rare que les tons ne se manifestent que dans les syllabes accentuées, et qu'un peu abusivement, on parle alors, par exemple, de deux « accents différents » selon que l'accent est accompagné d'un ton ou d'un autre.

On trouve, surtout en France, des comparatistes qui, du fait de l'absence dans leur langue d'un véritable accent, n'ont pas une connaissance intime de ce dont il s'agit. Inspirés par ce qu'ils savent du grec, ils refusent de parler d'un accent indo-européen, et font, en l'occurrence, constamment référence au « ton ». Or, il n'y a jamais, dans une langue, « un ton » au singulier. Le ton est une unité oppositive, ce qui veut dire qu'il y en a toujours plus d'un. Il y a certainement eu, au cours de l'évolution de l'indo-européen, apparition de tons : on en connaît en grec ancien, en balte, en serbo-croate, en scandinave. Mais il n'est pas du tout sûr que les plus anciens d'entre eux remontent à l'époque d'une communauté indo-européenne, même s'il y a, ici et là, un certain parallélisme.

Il y a quelques chances pour que des oppositions tonales se soient établies du fait de l'apparition successive de deux types distincts d'allongement de voyelle. Soit d'abord des voyelles allongées en compensation de la chute d'une consonne suivante (par exemple, un *-s* de nominatif). Le sommet mélodique se trouvera naturellement sur *le début* de la voyelle longue. Apparaissent ensuite des longues résultant de la coalescence, dans les flexions, de deux voyelles successives dont

la seconde était, au départ, accentuée, ce qui plaçait le sommet mélodique à *la fin* de la voyelle longue. Plus tard, lorsque de nouvelles longues ont résulté de l'allongement de voyelles par suite de la chute des « laryngales » suivantes, un nouveau type est apparu, avec sommet mélodique sur le début qui a pu refouler les anciennes du même type dans le camp des secondes, à sommet mélodique final. C'est ce qui semblerait attesté par les données baltes.

Rien n'empêche, bien entendu, qu'il y ait eu, à un stade très ancien, un système tonal en indo-européen, mais, si cela a été le cas, il n'en reste plus de traces perceptibles.

CHAPITRE X

LA GRAMMAIRE

La forme nue.

Dans des langues modernes comme le français ou l'anglais, il n'est, d'ordinaire, pas difficile de retrouver pour chaque substantif, chaque adjectif, voire chaque verbe, une *forme nue* non affublée de désinences : *loup*, sous sa forme écrite ou sous sa forme parlée /lu/, désigne l'animal bien connu sans impliquer aucune relation particulière avec les autres éléments de la phrase dans laquelle il peut figurer. La forme écrite peut se compliquer, au pluriel, par l'addition d'un *-s*, mais cela ne veut pas dire que la forme *loup*, sans *-s*, soit de ce fait marquée positivement comme un singulier : *loup* désigne l'espèce aussi bien que l'individu, comme on le voit dans *le loup a disparu des campagnes françaises*. En anglais, les choses se compliquent un peu : l'équivalent *wolf* peut prendre un *s*, dans *the wolf's*, avec la valeur de « du loup » ; il assume au pluriel la forme *wolves :* mais cela ne veut pas dire que *wolf* ne soit pas, comme *loup*, la forme nue du substantif sans addition sémantique aucune.

Si nous considérons maintenant l'équivalent latin *lupus*, nous constatons qu'outre la valeur « loup », la forme citée implique, de par sa finale, que s'il s'agit bien du loup, il doit être compris comme le sujet de la phrase. Comme objet, la forme serait *lupum*, comme complément du nom, *lupī*, comme complément d'attribution, *lupō*, etc. On serait tenté de dire que « loup » en latin se dit *lup-*. Mais cela serait inexact parce que « loup », c'est *lup-* plus un certain type de désinences. Soit, par exemple, une base latine *clāu- :* elle désignait, probablement, au départ un bâton, mais en fait, elle n'est attestée que sous la forme de mots différents : *clāua* « massue », *clāuos* « clou », *clāuis* « clef », chacun s'opposant aux autres par ses finales : au génitif, *clāua* fera *clāuae, clāuos* fera *clāuī, clāuis* restera *clāuis.*

Tout cela vaut pour les adjectifs. Sans restriction, en anglais, où « bon » est toujours *good*, invariable. En français, *bon* fera *bonne* au féminin, *bons* et *bonnes* au pluriel, mais, à tort ou à raison, *bon* /bõ/ est senti comme la forme de base correspondant exactement à la notion. En latin, l'adjectif *bonus* présentait les mêmes caractéristiques que le substantif *lupus*.

Pour les verbes, l'anglais ne pose pas de problème : la forme de l'infinitif, disons *say* « dire », s'identifie parfaitement à la notion, sans bavure sémantique aucune. En français, la forme de l'infinitif qui est censée représenter le verbe isolé, est, en fait, une forme marquée comme correspondant à des fonctions particulières. Mais, au moins dans la langue parlée, il existe une forme nue, celle de la troisième personne du singulier du présent de l'indicatif lorsqu'elle n'est pas accompagnée du pronom *il* ou *elle :* soit *chante, donne, mange* et, en faisant abstraction du *-t* de l'orthographe, *dit* /di/, *fait* /fɛ̀/, *court* /kur/. Rien de tel en latin où chaque forme verbale présente sa désinence personnelle et, éventuellement, ses affixes modaux : « chanter » c'est *can* + quelque chose qui varie de forme à forme, « dire » c'est *dīc* + *-o, -is, -it,* etc.

Dans les autres langues indo-européennes contemporaines, une forme nue des monèmes nominaux ou verbaux est souvent attestée : en russe *dom* « maison », *dobr* « bon » sont identifiables comme correspondant strictement aux notions correspondantes ; en allemand, on peut dire la même chose pour les équivalents *Haus* et *gut;* en espagnol *casa* « maison » est bien une forme nue à laquelle sera ajoutée le *-s* pluralisant, mais l'équivalent italien *casa,* avec son pluriel *case,* ne peut être considéré comme la base. Pour les verbes, l'allemand avec l'impératif, l'italien et l'espagnol avec leur troisième personne du présent de l'indicatif réalisent tant bien que mal une forme de base, encore que, par exemple, l'espagnol *quiere* « il aime » soit un candidat douteux lorsqu'on le rapproche de *quiero* « j'aime ».

Les amalgames.

Dans les langues anciennes de la famille, bien représentées par le latin, les formes nues paraissent absolument exceptionnelles ; en latin, par exemple, *dīc* « dis ! », *dūc* « conduis ! ». Partout triomphent des formes où radical et déterminations grammaticales sont amalgamées en autant de complexes pour lesquels on a trouvé la désignation de « mot ». Puisque, pour reconstruire l'indo-européen, on est parti de ces langues à amalgames, le résultat de la reconstruction ne pouvait être qu'une langue à amalgames, aussi longtemps qu'avec une louable probité, on ne se risquait pas au-delà de correspondances parfaitement régulières. Prenons un exemple. Soit l'accusatif pluriel du mot qui

désigne le cheval. Les données de la comparaison permettent de reconstruire **ekwons*. Si on a l'audace d'aller au-delà, on remplacera le *-n-* par un *-m-* sur le modèle de l'accusatif singulier et l'on interprétera le *-s* comme une marque de pluriel. Cela nous met sur la voie d'une analyse de la forme en une succession **eku* + *o* + *m* + *s* et d'une interprétation des trois derniers segments comme ce qui demeure d'anciennes particules plus ou moins autonomes auxquelles on pourrait attribuer respectivement les formes *-(H_1)e-*, *-(H_1)em-*, *-(H_1)es-*. Chacune de celles-ci représente, bien entendu, un faisceau d'hypothèses dont ceux qui osent les envisager ne se posent même pas la question de savoir si elles sont directement vérifiables, mais simplement si elles pourront finalement s'intégrer à une structure cohérente.

Dans les tout premiers temps de l'entreprise comparative, on a usé et abusé de ce genre d'hypothèses, et les chercheurs, d'esprit plus rassis, qui ont succédé aux pionniers, ont opportunément rappelé leurs disciples à un comportement moins désinvolte. Mais aujourd'hui, où nous avons une beaucoup plus vaste expérience des structures linguistiques et où nous sommes probablement moins impressionnés par la majesté des langues classiques, nous hésitons moins à dénoncer comme fonctionnellement détestables des formes où des unités de sens parfaitement distinctes reçoivent des signifiants variables selon les contextes et si intimement mêlées à d'autres qu'on ne peut savoir à quel point du discours elles se manifestent. La façon dont le système des déclinaisons latines s'est effondré dès que les sociétés héritières de Rome n'ont plus permis d'assurer le maintien des traditions linguistiques, porte témoignage de la foncière instabilité de structures linguistiques qui réclament pour être maîtrisées dans tous leurs détails l'appui d'un apprentissage qui se prolonge au-delà de la petite enfance.

C'est naturellement le jeu de l'analogie qui permet d'éliminer les complications qui résultent immanquablement de l'évolution phonique. Mais, dans un même domaine, l'analogie peut jouer dans des sens fort divergents : à partir d'*abstenir,* on pourra entendre un futur *abstenira,* mais à partir d'*abstiendra,* on a un infinitif *abstiendre ;* l'alternance isolée all. *bringen* « apporter », *gebracht* « apporté » sera éliminée par l'enfant, non en faveur de la forme la plus régulière qui donnerait le participe **gebringt,* mais sur le modèle d'autres verbes, irréguliers, mais fréquents, en *-ingen,* comme *singen* « chanter » qui fait au participe *gesungen,* d'où, en face de *bringen,* un participe *gebrungen.* On ne se dépêtre pas si facilement d'un bourbier où certaines circonstances, comme une particulière rapidité des évolutions phoniques ou la force des traditions, ont entassé des formes aberrantes.

Dans la recherche comparative, la méfiance vis-à-vis de l'analogie des formes était telle qu'on se refusait à rapprocher certaines désinences des formes adverbiales indépendantes, comme nous l'avons fait ci-dessus, p. 172, pour les désinences en $*b^h/*m$ et l'adverbe et fonctionnel *bei*, *by* de l'allemand et de l'anglais. Récemment[1], Jean Haudry a recherché en sanskrit les exemples d'éléments désinentiels non soumis à l'accord, c'est-à-dire manifestant une certaine autonomie vis-à-vis de leur radical, et reposé vigoureusement le problème de l'interprétation des désinences nominales à partir d'éléments autonomes, ce que nous allons retrouver ci-dessous.

LES NOMINAUX

Dans les langues indo-européennes, anciennes et modernes, on oppose nettement les verbes aux nominaux. Les verbes sont par nature des noyaux de proposition entourés de satellites grammaticaux qui précisent la nature du procès (action ou état) exprimé, certaines circonstances de ce procès et les rapports à ce procès des participants à l'acte de communication. Les nominaux, noms, pronoms, adjectifs, peuvent également fonctionner comme noyaux de propositions, dans ce qu'on appelle les propositions nominales : lat. *Paulus bonus* « Paul est bon », russe *dom nov* « la maison est neuve », mais leurs satellites grammaticaux sont en priorité, moins des déterminations qui apportent des précisions sur leur valeur, comme, par exemple, le pluriel, que des indicateurs de leurs rapports avec le contexte dans l'énoncé. Ces indicateurs de leur fonction sont, dans une langue très évoluée comme le français, des éléments formellement autonomes, les prépositions, mais, dans les langues anciennes, des appendices amalgamés qu'on appelle les cas.

Les nominaux comportent les noms proprement dits qu'on désigne aussi comme les substantifs, les pronoms qui désignent, comme les noms, des entités variées, mais avec une moindre spécificité, et les adjectifs censés marquer certains aspects de ces entités. Un des traits qui caractérisent l'indo-européen à date ancienne par rapport à d'autres familles linguistiques est le fait que les adjectifs se combinent avec les mêmes satellites grammaticaux que les substantifs, alors qu'ailleurs, ils se comportent plutôt comme des verbes : en latin, « blanc » se dit *albus* qui se « décline » comme le nom *dominus* « maître », alors que, dans beaucoup de langues, on emploiera un verbe avec la valeur d' « être blanc ». Il y a en latin des traces de cet

1. Dans *Préhistoire de la flexion nominale indo-européenne*, Institut d'études indo-européennes de l'Université Jean Moulin (Lyon III), 1982.

usage : à côté d'*albus*, on a un verbe *albet* « il est blanc ». Il y a même en français une trace de ce genre de forme dans les verbes en *-oyer* comme *verdoyer*, connu surtout sous la forme de son participe présent *verdoyant*. Mais les usagers contemporains ont du mal à comprendre *verdoyer* comme correspondant à un état et sont tentés de l'interpréter, non comme « être vert », mais comme une sorte d'équivalent de *verdir*, c'est-à-dire « devenir vert ». C'est bien pourquoi ce verbe ne s'entend plus que dans des récits traditionnels où le sens précis n'a plus grande importance.

En ce qui concerne le statut grammatical des nominaux, la reconstruction a longtemps visé essentiellement à restituer des déclinaisons primitives, c'est-à-dire les formes qu'on pensait pouvoir affecter aux complexes formés des radicaux et des différents cas amalgamés. On se demandait moins quelle avait pu être la valeur de telle ou telle désinence attestée dans une ou plusieurs langues de la famille, que lesquelles, parmi les formes attestées, devaient figurer dans telle ou telle des cases posées au départ, en s'inspirant du système le plus « riche », celui du sanskrit. En d'autres termes, on imaginait un système fermé sur le modèle de celui qu'on avait établi pour classer les formes du latin et du grec, plutôt qu'un état de langue où étaient en compétition diverses formes, plus ou moins isolables.

Dès qu'on est convaincu que la reconstruction n'a de sens que dynamique, c'est-à-dire qu'elle doit viser à retrouver les grands traits d'une évolution, on ne saurait se satisfaire d'un cadre préétabli et il convient de partir des formes attestées, de chercher à leur attribuer une valeur à tel point de l'évolution, c'est-à-dire en opposition avec telle ou telle autre existant à la même époque, de rapprocher celles qui semblent dérivées d'une même forme ancienne et de tenter d'expliquer le conditionnement des divergences formelles.

A la recherche des radicaux nus.

Un premier effort doit viser à retrouver, dans les langues attestées, les radicaux nus, ceux qui ne présentent pas trace d'amalgame de désinences casuelles ou de déterminations grammaticales. De telles formes ne sont pas rares, mais, comme nous l'avons vu, elles n'existent guère, dans les langues classiques, que comme variantes un peu marginales de complexes à éléments grammaticaux amalgamés : une forme latine comme *animal* « animal », qui ne comporte rien qui indique une fonction quelconque dans la proposition où elle figure, est interprétée, non comme le monème *animal* libre de toute implication grammaticale, mais comme le cas sujet et le cas objet indifférenciés, donc comprise en fonction de son intégration à une déclinaison

comportant un génitif *animalis*, un datif *animali*, un nominatif-accusatif pluriel *animalia*, etc. Cette vision des faits n'est d'ailleurs pas limitée aux langues anciennes, elle conditionne notre conception des faits contemporains : pour un Français qui a été sur les bancs de l'école, *loup* n'est pas l'équivalent de la notion correspondante, mais cette notion plus celle de singulier.

Les radicaux nus les plus fréquents sont certainement les formes du type *animal*, c'est-à-dire les neutres autres que ceux qui se terminent en *e/o*, à savoir les radicaux en consonne, comme lat. *animal* lui-même ou lat. *caput* « tête », en *-u* comme lat. *pecu* « le petit bétail », ou en *-i*, comme lat. *mare* (< **mari*). La meilleure définition des neutres est celle qui les caractérise comme ne distinguant jamais la forme du sujet et celle de l'objet. On peut alors légitimement se demander comment les usagers de la langue vont s'y retrouver et l'on reviendra abondamment sur ce problème un peu plus loin.

D'autres radicaux nus ont des chances de se retrouver parmi les formes dites de vocatif, celles qui servent à appeler. En fait, les formes de vocatif distinctes de celle du cas sujet, dit nominatif, et qui doivent correspondre à des formes anciennes, ne sont guère attestées que pour les noms qui se terminent en *e/o*. En latin, par exemple, le vocatif de *dominus* « maître » est *domine*, forme nue d'un dérivé formé à partir de *dom-* « maison » au moyen des deux suffixes *-e/on* + *e/o*. Ce vocatif est largement attesté à date ancienne et on l'utilise encore régulièrement en tchèque : une personne prénommée *Roman* sera interpellée comme *Romane !* Naturellement, ce vocatif tchèque n'est plus une forme nue, depuis que l'usure des finales a éliminé les désinences de sujet et d'objet des formes masculines. On peut dire que le vocatif ne l'était déjà plus en latin où la partie de *dominus* non affectée par les finales était *domin-* et non plus, comme à date très ancienne, **dom-en-o-* auquel s'*ajoutaient* les désinences casuelles.

D'autres vocatifs remontent à des formes du radical où la voyelle finale a été abrégée. C'est le cas des vocatifs en *-a* de substantifs en *-ā* qui sont attestés en slave sous la forme attendue *-o*, en tchèque, par exemple, *ženo* « femme ! ».

Les radicaux nus se rencontrent comme premiers éléments de composés, par exemple, dans le *dēmo-* de grec *dēmokratía* « démocratie » ou dans l'*agri-* de lat. *agricola* « agriculture » régulièrement dérivé d'**agro-* « champ » qu'on trouve intact dans grec *agronómos* « agronome ». Ceci laisse supposer qu'anciennement, les rapports marqués plus tard par le cas dit « génitif », ceux qui s'expriment en français contemporain au moyen de la préposition *de* entre deux substantifs, résultaient simplement du rapprochement des deux termes, vraisemblablement en faisant précéder le noyau par son déterminant. Le vieux-français a connu un état de chose analogue, mais avec postposi-

tion du déterminant. Il nous en reste *Hôtel-Dieu* et le modèle productif représenté par *rue Gambetta, avenue Sadi-Carnot.*

D'un intérêt tout particulier est la forme nue des radicaux de certains pronoms, *tu* « tu », comme exemple de personnel, et **so,* comme démonstratif, qui a donné l'article défini du grec, *ho*. On y reviendra longuement ci-dessous.

On dispose aujourd'hui du terme *absolutif* pour désigner une telle forme nue lorsqu'elle a un statut syntaxique particulier. Ceci n'est pas le cas, dans les langues indo-européennes attestées, où de telles formes sont interprétées comme des « nominatifs » au même titre que les formes en *-s* que nous allons trouver ci-après. Nous ne les désignerons comme des *absolutifs* que dans des états de langue que nous poserons antérieurement à leur confusion syntaxique avec les formes en *-s*. On peut penser que tout substantif pouvait, très anciennement, apparaître sous la forme d'un absolutif, non seulement dans l'appel (vocatif), dans la nomination (nominatif proprement dit) et dans des emplois de valeur génitivale, mais, comme nous le verrons plus loin, en référence au participant unique (« sujet ») d'un procès intransitif et au patient (« objet ») d'un procès transitif.

Ablatif et allatif.

Pour marquer qu'un participant d'un procès avait un rôle actif, on disposait d'une particule qu'on reconstruit comme **es* (avec ou sans « laryngale » H_1 préposée)[2]. Il est probable que, dans un premier temps, l'emploi de cet **es* était facultatif, c'est-à-dire limité aux cas où l'on pouvait hésiter sur l'identité de l'agent et du patient : « homme » + « tigre » + « action de tuer » (qui tue l'autre ?). Cette particule marquait en fait l'éloignement « à partir de » : l'agent était celui d'où provenait l'action ; elle était également la marque du point de départ, ce qu'on appelle plus tard l'ablatif, et aussi celle de l'origine, ce qui aboutira, un jour, au génitif. On peut poser, comme valeur initiale de cet élément, celle de l'anglais *from*.

S'opposait à cette particule une autre, reconstruite comme **ey* (ou H_1*ey ?*), marquant le rapprochement. On peut lui accorder les valeurs de l'angl. *to*. Ce rapprochement, dans l'espace, correspondait à un déplacement dans la direction de l'entité désignée, « à » dans *il va à Paris*, à l'exclusion de la relation marquée par *à* dans *il est à Paris.* Lorsque l'entité désignée était un être animé, **ey* pouvait marquer

2. Une première esquisse de l'exposé qui suit figurait dans un compte rendu, par l'auteur de ces lignes, du livre de T. BURROW, *The Sanskrit Language*, dans *Word* 12, 1956, p. 304-312, reproduit dans *Evolution des langues et reconstruction*, Paris, 1975, p. 99-105.

l'attribution : « donner à quelqu'un ». Ces deux valeurs peuvent être désignées, la première comme l'allatif, la seconde comme le datif.

Le locatif.

On peut penser que ce qu'on appelle le locatif simple, l'endroit où a lieu le procès, marqué en anglais par *at*, s'exprimait normalement par la simple désignation du lieu, comme on le fait, aujourd'hui encore, en français, dans un énoncé comme *je l'ai rencontré place de l'Opéra*, ou peut-être déjà au moyen de **en*, qui a donné le français *en*, lorsque le lieu était considéré non comme un point, mais comme un espace cerné : *le sport en chambre*. Toutefois il est constant que la désignation de l'endroit où l'on va finisse par s'employer également pour l'endroit où l'on est, comme c'est le cas en français et comme le montre l'évolution de **ad* qui avait au départ, encore en latin, une valeur allative, mais qui en anglais, sous la forme *at*, est un strict locatif sans déplacement. On peut donc supposer que cette tendance s'est très tôt manifestée dans le cas de **ey*.

Ces deux particules **es* et **ey*, dont on peut représenter graphiquement la valeur comme •⟶ et ⟶•, ont, respectivement, à l'époque où les voyelles ne se sont maintenues que sous l'accent, été traitées de façons différentes, essentiellement en fonction des valeurs particulières qu'elles pouvaient assumer. Normalement placées après le substantif dont elles explicitaient la fonction, elles ont, dans certains cas, été accentuellement intégrées à ce substantif et, comme c'était alors la norme dans ce cas, elles ont porté l'accent et gardé leur voyelle, le résultat étant *-és*, *-éy*. Dans d'autres cas, restées hors du cadre accentuel du substantif précédent, elles ont perdu leur voyelle et ont abouti respectivement à *-s* et à *-i* qui, finalement, se sont agglutinés au substantif précédent.

La forme accentuée en *-és* est celle qui s'est finalement imposée pour le complément du nom, le génitif. Elle entraînait régulièrement la chute de la dernière voyelle du radical, puisque c'est *-és* qui portait l'accent du mot : soit, par exemple, *ped-* « pied » ; au génitif la forme attendue est **pdés* dont on trouve encore des traces, mais où les locuteurs ont en général rétabli, par analogie, la voyelle du radical, d'où *pedés* qui est à la source de la forme latine *pedis*. La forme *-és*, à voyelle pleine, se rencontre également avec une valeur spatiale, c'est-à-dire comme ablatif, avec chute, puis restitution de la dernière voyelle du radical.

La forme en **-éy* est celle qui s'est assez largement établie comme la marque du datif avec, ici aussi, une forme attendue **pdéy* où, comme précédemment, la voyelle du radical a été généralement rétablie par analogie. Mais il n'est pas rare qu'on rencontre aussi au datif, non plus

-éy, mais la forme inaccentuée de la désinence, donc *-i*, comme dans le grec *podi*.

Le sort du -s de nominatif.

La particule *s*, finalement agglutinée comme *-s* au radical, s'est établie comme la marque générale de l'agent du procès, c'est-à-dire comme un ergatif. Là où le radical se terminait sur la voyelle *e/o* l'*s* s'est très généralement maintenu dans les langues les plus anciennement attestées, même lorsque la forme qu'il caractérisait a eu acquis le statut de nominatif, c'est-à-dire de forme conçue hors-syntaxe. C'est la source de tous les mots en *-os* du grec, en *-us* du latin, en *-as* du sanskrit. Il en allait de même pour les radicaux en *-eu-*, *-ei-* comme *H_3ewei-* « mouton » qui perdaient leur deuxième voyelle devant *-s*, d'où *H_3ewi-s*, lat. *ouis*. Là où le radical se terminait sur une consonne occlusive, le *-s* a probablement été éliminé à maintes reprises, au moins dans certains contextes comme, par ex., lorsqu'un *-t* + *s* final était suivi d'un mot commençant par *s-*, *-ts s-* passant alors à *-t s-*; le groupe *-ts* lui-même a dû fréquemment aboutir à *-s* pour se voir restitué par analogie par les générations suivantes. En latin, par exemple, on en est à un stade *-s* dans *mīles* « soldat » nominatif, en face de l'accusatif *mīlitem* qui conserve le *-t-* de la base.

C'est lorsque la base se terminait sur une consonne continue, liquide, nasale ou « laryngale », que le sort de cet *s* pose surtout des problèmes. Aussi longtemps qu'il restait la marque d'un cas d'agent, de fonction syntaxique bien déterminée, il a dû être, là aussi, très régulièrement restitué toutes les fois où une évolution phonétique tendait à l'éliminer. Mais lorsque le cas en *-s* est devenu un « nominatif », c'est-à-dire une forme utilisable hors syntaxe, il a pu être ressenti comme assez naturel qu'il se confonde avec le radical nu, et les locuteurs ont été moins tentés de restituer le *-s*. C'est ce qui nous vaut des formes en *-r* comme grec *patēr*, lat. *pater*, got. *fadar*, où il n'y a pas trace de *-s*. En sanskrit, l'équivalent *pitā* avec chute du *-r* peut être la trace d'un processus général de chute du *r* final ancien, peut-être sous la pression du passage de *-n* à *-r*[3]. Dans ce cas, le *-r* des autres langues serait restitué analogiquement à partir des autres cas. Cette disparition de *-r* pourrait être en rapport avec l'allongement de la voyelle qui, ici comme ailleurs, dans les radicaux à finale consonantique, caractérise le nominatif. Les fréquents passages de *s* à *r* qu'on relève, par exemple, en latin (*generis* pour *genes-es*, de *genus* « genre ») et en germanique (pluriels scandinaves en *-r* en face du *-s* de l'anglais : dan. *træ-er* « arbres », angl. *tree-s*) portent témoignage de l'apparente-

3. Voir, ci-dessus, p. 171.

ment articulatoire des deux phonèmes. Un *-r-s* final était constamment exposé à donner un *-r* (articulé comme le *s* d'alors, avec la pointe de la langue relevée) que de nouvelles générations confondaient avec le *-r* ordinaire.

Le mot pour « sel » illustre les traitements possibles d'un radical en *-l* : le grec *hal-s* présente le *-s* et la voyelle normale du radical ; le latin *sāl* a perdu le *-s*, mais présente en revanche l'allongement de la voyelle.

Dans les radicaux en *-s*, la finale de nominatif *-s-s* a été régulièrement réduite à *-s* avec allongement de la voyelle précédente, comme l'illustre, par exemple, le grec *eumenḗs* « bienveillant » au masculin-féminin, de **eumenes-s*, en face du radical nu du neutre *eumenés*.

Sur le sort de *-s* après les radicaux en nasale, la comparaison des formes grecques et sanskrites du terme qui désigne la terre illustre bien les vicissitudes auxquelles il a été exposé. Le grec a au nominatif *khthṓn* sans *-s* et une voyelle longue, en face du génitif *khthon-ós*. Le sanskrit a le nominatif *kṣā́s* avec *-s*, voyelle longue et chute de la nasale, en face du locatif *kṣámi* et du génitif *jmas* qui attestent la nasale *-m-*. Dans ce mot très fréquent, l'analogie a peu joué en sanskrit, l'enfant apprenant très tôt les diverses formes résultant régulièrement de l'évolution phonétique. En laissant de côté le consonantisme initial qui, nous l'avons vu ci-dessus, p. 167, soulève des problèmes, on posera un radical en *-om-*. Le nominatif *-om-s* peut évoluer en *-on-s*, voire en *-ōs* avec nasalisation de la voyelle. C'est là, sans doute, ce qui a donné la forme sanskrite après dénasalisation de la voyelle. A partir de *-on-s*, l'analogie peut étendre le *n* aux autres formes du mot, d'où le génitif *khthon-ós* ; le stade ultérieur *-ōs* > *-ōs* fournit la voyelle longue du nominatif, mais l'analogie va y réintroduire le *-n-*, d'où *-ōns* qui se simplifie en *-ōn*. Le latin présente le *-m-* partout dans les formes apparentées *humus* « terre », *homo* « homme » (c'est-à-dire « le terrestre », par opposition aux dieux).

Le cas des radicaux en « laryngales » est le plus complexe. Nous avons vu qu'on peut, au moins pour certaines d'entre elles, poser qu'elles se sont durcies en *-k-* devant l'*s* du nominatif. Mais, bien entendu, il n'est nullement acquis que l'analogie va faire triompher la forme en *-k-* aux dépens de celle des autres cas. A partir de radicaux en *-H_2*, par exemple, on peut s'attendre à des nominatifs en *-ās* (grec *neaniā-s* « jeune homme », en *-aks* (lat. *senex* < **senak-s*) ou analogiquement en *-āks* (lat. *audāx* « audacieux »). Mais il faut également envisager que, pour un complexe **-eH_2-s*, toujours analogiquement restituable à partir des autres cas, le traitement normal, à certaines époques, a pu être *-ās*, ou, avec chute de *-s* après une continue H_2, *-ā*. C'est cette dernière forme qu'on rencontre dans les noms masculins du type lat. *agricola*, russe *voevoda*, en face des *-ās* (ou *-ēs*<*-ās*) du grec.

Le cas des féminins est tout particulier. Il faut là distinguer entre le suffixe de sexe et la marque du genre.

On a vu ci-dessus, p. 157 que le suffixe traditionnel pour marquer le sexe féminin peut être reconstruit comme *-ye/oH₂-* et, avec vocalisme zéro, *-iH₂-*. Devant *-s*, ce dernier peut aboutir à *-ik-* et analogiquement *-īk-*, *-īs* ou *-ī*, parallèlement à ce qu'on constate pour les masculins qui précèdent. H_2, ici, doit être le même que celui qu'on pose pour le **-ā* dont on trouve le reflet dans *senex, audāx, agricola, voevoda* ou *neaníās*.

Le démonstratif.

En revanche, le H_2 que l'on peut reconstruire pour la marque du genre féminin, a une tout autre origine. Il s'agit d'un élément qui a dû s'agglutiner au pronom démonstratif lorsqu'il servait à désigner une femme. Il faut s'attarder un peu sur ce pronom démonstratif qui soulève des problèmes divers. Il semble qu'il était strictement au départ ce qu'on appelle un déictique, c'est-à-dire qu'il avait à peu près la valeur du geste qui consiste à montrer du doigt. Lorsque le geste se fait, non plus avec le doigt, mais avec la langue, c'est la pointe de cet organe qui entre en action. Le produit phonique est une apicale, un [t] suivi d'un [a] ou d'une voyelle d'avant. Un peu partout, les enfants paraissent retrouver ce démonstratif naturel qui est interprété par les adultes comme une déformation d'un des mots traditionnels de la langue, en français *tiens!* en italien *dai* ou *togli* « prends », en anglais *there* « là », en allemand *da* « là », etc. Une fermière qui donne du grain à sa volaille ponctue volontiers son mouvement d'un *ta, ta, ta!*

Cette production phonique naturelle, fréquemment répétée, comme on le voit chez la fermière, est certainement à la source du déictique indo-européen. On attendrait donc **tata*, ou avec un vocalisme évolué, **tete* ou **toto*. Lorsque intervient l'accent qui aboutit à faire tomber celles des voyelles qui peuvent disparaître sans dommage, il semble qu'avec la valeur du français *ça!* l'accent ait porté sur la première syllabe, d'où **tóto* puis **tot*. Cet accent non final est, d'ailleurs, caractéristique des noms neutre. Au contraire, pour désigner des personnes, l'accent portait sur la finale, donc **totó* et, avec chute de la voyelle accentuée, **tto*. Or, on l'a vu, lorsqu'au cours de l'évolution phonique ou du fait de la dérivation, deux *t* se sont trouvés en contact, ils se sont réduits à *s-*. On a donc abouti aux deux formes **tot* et **so* largement attestées, en gotique, par exemple, sous la forme des démonstratifs *þat (a)* et *sa*, en grec, avec valeur d'article défini, *to* et *ho*. L'article et le démonstratif de l'anglais et de l'allemand, *the, der, die, das, that, dass*, remontent à ces formes, avec extension analogique du *-t*. Un phénomène tout à fait analogue s'est produit en russe

ancien à partir du **to* qui avait analogiquement remplacé **so* : le masculin *tot* y provient d'une forme redoublée **tŭtŭ*.

Au départ, il s'agissait strictement d'une *forme hors-syntaxe*, donc invariable. Lorsque, ultérieurement, le terme s'est généralisé dans d'autres emplois et a reçu des désinences casuelles, c'est la forme de base **to* qui a servi, comme on le voit par l'article du grec : nominatif grec *ho* (< *so*), mais accusatif *tón*, génitif *toû*, datif *tôi*.

L'apparition d'une forme plurielle de ce déictique n'a pas coïncidé avec celle du pluriel des substantifs. Cette forme, **toi*, pourrait résulter de l'adjonction à **to* de la particule *i* indicative de la présence *hic et nunc*, qu'on retrouve largement comme marque du présent vrai, et en grec comme renforcement des démonstratifs, par exemple, dans *houtos-í* « celui-ci » avec insistance sur « -ci ». On peut comprendre comment une forme d'insistance a pu être sentie comme indication de pluralité.

Le genre féminin.

L'existence de trois formes distinctes, **so*, **to*, **toi*, pour référer à des situations différentes : être animé, être inanimé et circonstance « pluralité », a pu suggérer la création d'un quatrième déictique pour référer aux femmes et aux femelles. Le H_2 qui a servi à cette fin s'est naturellement agglutiné à **so*, désignation des animés. On peut penser qu'il rappelle le -H_2 final du radical qui désigne la femme : grec *gunḗ* (cf. *gynécologie*), russe *žena*, gotique *qino*, anglais ancien *quean* (« femme de mauvaise vie »), v. irl. *ben ;* avec une voyelle longue et une autre flexion, on a got. *qēns*, anglais *queen*. Ce radical se reconstruit comme $*g^{wo}neH_2$, et le démonstratif féminin seH_2 a pu en être conçu comme un rappel sous forme de rime.

Il est important de ne pas oublier que **so* et son féminin seH_2 sont restés longtemps des formes hors-syntaxe, en quelque sorte des gestes dissociés des énoncés organisés, précédant ou suivant. Ces formes, d'une extrême fréquence, ont pu, de ce fait, ne pas céder aux analogies et conserver, à travers les siècles, leur forme nue, sans adjonction de *-s* de nominatif. C'est le sentiment persistant de l'identité du -H_2 de seH_2 et de celui des formes féminines en -H_2 qui *a toujours empêché l'extension du -s de nominatif aux féminins en -H_2* et à leurs successeurs en *-ā*, alors que les masculins de même forme étaient toujours exposés à recevoir cette désinence.

L'existence d'une différence entre **so* et seH_2 n'impliquait nullement, en elle-même, celle d'un genre féminin distinct d'un genre masculin, pas plus que la distinction en français entre *ça* et *celui-là* n'implique l'existence d'un genre neutre. Le genre commence à partir du moment où il y a **accord**, c'est-à-dire redondance. Il n'y a plus de

genre en anglais parce que la référence à certains êtres humains, voire à certains moyens de locomotion, bateaux ou voitures, au moyen de *she* au lieu de *he* ou de *it*, apporte un élément d'information. Ce n'est pas un « accord ». La même chose vaut en français lorsqu'on dit *le docteur... elle*, où *elle* nous fait savoir que le docteur est une femme, ce qu'on pouvait ignorer. Mais si l'on entend *la chaise... elle* dans un contexte où *elle* renvoie évidemment à *chaise*, l'information ajoutée par *elle* au lieu de *il* est nulle. Le genre féminin de *chaise* qui se manifeste par des accords, celui de l'article, *la*, et celui des adjectifs, *petite*, *grande*, ne contribue en rien à la communication de l'expérience, parce que tous les usagers savent qu'en rapport avec *chaise*, il *faut* dire *la*, *petite*, *grande*. En indo-européen, le genre féminin est apparu lorsque la forme seH_2, uniquement pronominale au départ, s'est vue transférée aux emplois adjectivaux : aussi longtemps que « cette femme », où le démonstratif est en présence du substantif qu'il détermine, a eu la forme $*so\ g^{we}neH_2$, sans « accord » de *so*, en face de $*g^{we}neH_2$... ... seH_2, il n'y avait aucune redondance, c'est-à-dire aucun genre féminin. Mais lorsque, sur le modèle de l'emploi pronominal, s'est imposé $*seH_2\ g^{we}neH_2$, le ver était dans le fruit. De proche en proche, s'est imposée la nécessité de choisir entre **so* et $*seH_2$ comme l'adjectif démonstratif correspondant à chaque substantif et finalement l'obligation d'employer l'un OU l'autre pour chaque substantif.

Le cas des démonstratifs anglais illustre bien comment un accord peut apparaître lorsqu'un pronom s'emploie comme adjectif. En anglais, comme on le sait, le pluriel ne s'exprime que dans les substantifs, sans que les adjectifs, épithètes ou attributs, en soient affectés : *a nice girl* « une jeune fille aimable », au pluriel *nice girls*, *the girl is nice* « la jeune fille est aimable », *the girls are nice* « les jeunes filles sont aimables », sans trace d'une marque de pluriel dans *nice*. On devrait donc avoir en face de *this girl* « cette jeune fille-ci », *that girl* « cette jeune fille-là », les pluriels **this girls*, **that girls*. En fait, on a *these girls*, *those girls* où l'on retrouve *these* et *those*, pluriels des pronoms *this* et *that*. Cette entorse à la « règle » du non-accord de l'adjectif reste, heureusement, limitée à ce cas particulier.

En indo-européen, l'emploi de $-H_2$ pour indiquer la référence aux personnes du sexe féminin a pu s'étendre à d'autres pronoms démonstratifs dans la mesure où leur forme se prêtait à cette addition. Une fois établi pour seH_2, l'accord a dû s'étendre à ces autres démonstratifs. L'alternance $-o/-eH_2$ (êtres non féminins/êtres féminins) a dû vite s'imposer à d'autres pronoms où elle apportait une intéressante information : « celle-là », au lieu de « celui-là », avec, comme ci-dessus, extension aux emplois adjectivaux des dits pronoms. Elle a finalement gagné les adjectifs en *-o*, peut-être d'abord dans leurs emplois avec valeur de substantif : **newo-* « le neuf » > $*neweH_2$ « la

neuve », puis avec généralisation aux cas d'emploi prédicatif ou épithétique.

Lorsque tout adjectif ou emploi adjectival en *-o* s'est trouvé affublé d'un féminin en *-eH₂*, s'est posée pour tout substantif la question de savoir laquelle des deux formes devait être employée. Soit **snuso-* « la bru ». Devait-on se laisser influencer par la forme et dire **so snuso, newo- snuso-*, ou plutôt, ce qu'on constate en fait, utiliser les formes en $-eH_2$, donc $*seH_2$ *snuso-*, $*neweH_2$ *snuso*, sur le modèle de **snuso-*... ...seH_2, **snuso-*... ...$neweH_2$ où ces formes apportaient une information ?

La comparaison des langues dans lesquelles s'est établie la distinction du féminin et du masculin nous révèle un état où seuls les adjectifs en *-o* étaient affectés par l'alternance. En d'autres termes, les radicaux en *-i*, *-u* et consonnes n'étaient pas affectés par l'accord : en latin, *grandis* « grand » a les mêmes formes, qu'il se rapporte à un nom féminin ou à un nom masculin. C'est au cours de l'histoire du français que s'est établie l'alternance *grand* - *grande*, sur le modèle de *bon* - *bonne*, l'absence d'accord étant conservée dans des formes figées comme *grand mère* ou *grand rue*.

Toutefois les adjectifs en *-o* étaient très nombreux et, pouvant se combiner avec n'importe lequel des substantifs, ils imposaient, pour chacun de ces derniers, le choix de *-o* ou de $-eH_2$, c'est-à-dire un genre, le féminin ou le masculin. Ce choix a dû être déterminé soit par la forme du substantif, soit par sa valeur. Nous avons vu que là où le radical du mot *langue* se terminait en $-eH_2$, le mot était féminin, comme en latin ou en germanique. Là où, le *-e-* étant tombé, on était parti de $-uH_2$, le mot était masculin, comme en balte ou en slave. Par ailleurs, un mot comme **snuso-*, qui désignait un être féminin, a fait des accords « féminins », c'est-à-dire a pris le « genre féminin ». De même, à date ancienne, les noms d'arbres étaient du genre féminin, sans doute parce qu'ils étaient perçus comme des êtres femelles qui donnaient naissance aux fruits : en latin, par exemple, *fīcus* « figuier », bien que présentant un ancien radical en *-o-*, détermine des accords féminins, c'est-à-dire *est* du genre féminin.

Pendant longtemps cet $-eH_2$, ou son équivalent plus récent *-ā*, n'a servi que comme marque du genre, c'est-à-dire limité aux pronoms et aux adjectifs. Il y avait certes des noms à radical en $-eH_2$ qui tendaient à être du genre féminin sauf lorsqu'ils désignaient des êtres de sexe masculin. Mais $-eH_2$ n'était nullement un suffixe de dérivation pour former des désignations de femelles. En latin, la louve reste longtemps *lupus fēmina* et l'agnelle *agnus fēmina*. Le suffixe de sexe féminin est, on l'a vu, tout autre. C'est $*yeH_2$, attesté, le plus souvent, sous la forme *-ī-* ou *-īk-*. L'utilisation de *-ā* ($< eH_2$) comme suffixe substantival de sexe féminin a dû être longtemps freinée par l'existence du

suffixe homonyme de singularisation d'*agricola*, *voevoda*. Il n'est guère possible de se prononcer sur la question de savoir si cette homonymie existait de fondation ou si ce que nous notons uniformément *eH₂* et, plus tard, *-ā* présentait à l'origine deux « laryngales » distinctes (/χ/~/ɣ/ par exemple).

Le fait que la distinction d'un démonstratif féminin se soit réalisée, au départ, dans des emplois asyntaxiques a eu pour effet constant de bloquer toute extension analogique du *-s* de nominatif aux féminins en *-ā*. Lorsque aurait pu se produire ce phénomène, il existait certainement beaucoup de radicaux en liquide, nasale ou « laryngale » où l'évolution phonétique avait abouti à éliminer le *-s* de nominatif, même si l'analogie pouvait amener de nouvelles générations à le restituer régulièrement ou épisodiquement : à côté de laryngales durcies en *-k-* devant le *-s* de nominatif et qui, finalement, en assuraient la permanence, il y avait probablement une majorité de cas où la « laryngale » avait, en quelque sorte, absorbé cet *-s*. Les nominatifs en *-eH₂* (ou déjà *-ā*) pré-existaient sans doute à l'apparition du *-eH₂* de *genre* féminin. Mais alors que les masculins en *-eH₂* (*-ā*) ont pu, à tout moment, récupérer un *-s* analogique comme marque de nominatif, cela ne s'est jamais produit pour les féminins où se maintenait la tradition de l'emploi asyntaxique originel.

Les substantifs en *-o*, quant à eux, présentaient, de fondation, des nominatifs en *-s* et même si, au cours de l'évolution de la langue, leur flexion a été constamment influencée par celle du démonstratif, il n'était pas question que la pression de l'asyntaxique *so* puisse déloger cet *-s*.

Le génitif.

La formation du génitif au moyen de la particule *es* s'est réalisée le plus souvent sans les complications qu'on relève pour le nominatif où, comme on vient de le voir, la combinaison de *-s* avec une consonne précédente a créé des problèmes. On relève cependant que, lorsque l'accent de la forme nue du radical ne se trouvait pas sur la dernière syllabe, l'adjonction du *es* du génitif a eu pour effet, le plus souvent, d'attirer l'accent sur la dernière syllabe du radical et non sur *es*. Soit, par exemple, un des mots qui désignent le feu, qu'on peut reconstruire comme **H₁égni-*. L'accent s'y trouvait sur la première syllabe. Le *-s* du nominatif s'est ajouté sans problème au radical nu, d'où skrt. *agni-s*, lat. *igni-s* ; au génitif, l'accent portait sur la seconde syllabe, ce qui a donné **H₁gnéi-s*, mais, avec restitution analogique de la première voyelle, **H₁egnéi-s*, skrt. *agné-s* : on y constate la chute de la voyelle de *es*. Le maintien de la voyelle, qui laisse supposer une ancienne accentuation, est attesté, par exemple, pour **H₃ewi* « mouton », skrt.

avi, avec le génitif *avyas* < *H_3ewyés*. Mais le traitement précédent est plus généralement attesté.

Il y a eu, d'autre part, des cas où, du fait de diverses circonstances, la forme du génitif et celle du nominatif risquaient de se confondre. Il semble que ceci ait été le cas pour les radicaux en voyelle. On peut, par exemple, reconstruire pour le loup, le nominatif **wl̥kʷé-s*. Le génitif doit être **wl̥kʷ-és* avec chute de la voyelle du radical. On en est encore à l'époque où la voyelle inaccentuée disparaît. Ce n'est que plus tard que, l'accent étant moins fort, deux voyelles en contact dans la flexion ou la dérivation vont donner une voyelle longue. Le résultat va donc être une forme **wl̥kʷés* identique pour les deux cas. C'est bien ce qui semble être le cas en hittite où *ata-s* correspond aussi bien à « père » nominatif qu'à « du père » génétif. Mais il n'est pas exclu qu'il y ait eu entre les deux quelque différence que la graphie ne fait pas apparaître. Lorsque la forme en *-s*, devenue un nominatif, voit son emploi multiplié, la nécessité de la distinguer du génitif en *-es* devient sans doute pressante et c'est pourquoi on constate que toutes les langues traditionnelles ont trouvé une solution au problème. La plus répandue est celle qui consiste à emprunter le génitif du démonstratif. C'est le cas en germanique, par exemple, où *dags* « jour » présente, avec *-s*, le traitement régulier d'un nominatif en *-es*, alors que *dagis* « du jour » rime avec le génitif *þis* du démonstratif, représentant sans doute un ancien **teso*, c'est-à-dire *t-es-* plus un *-o* adjectival. C'est le cas en sanskrit où, en face de *dev-a-s* « dieu », on a *dev-as-ya* « du dieu », avec un **-yo* adjectival ajouté à **deiw-es-*. En grec, on a hésité entre *-o* et *-yo*, d'où, en face de *hippo-s* « cheval », on a *hippoio (*hippos-yo)* ou *hippou* (< **hippos-o*). Le slave s'est tiré d'affaire au moyen du nouvel ablatif résultant de l'adjonction au radical en *-o* de l'adverbe **od* de provenance (russe *ot* « à partir de ») : v. slave *rabŭ* « esclave » (< **orbo-s*), génitif *raba* (< **orbōd* < *orbo-od*). En latin et en celtique, on a eu recours à une forme toute différente, celle d'adjectifs dérivés de radicaux à consonne au moyen du suffixe **-y-eH* > **-iH*, d'où *-ī*. On a donc, en face de lat. *dominu-s* « maître », *domin-ī* « du maître » ; en v.-irlandais *fer* « homme » vient de **wiro-s*, *fir* « de l'homme » de **wir-ī*.

Le génitif et l'**ablatif** sont restés longtemps indifférenciés, l'un et l'autre indiquant l'origine. On vient de voir, par le slave, que l'ablatif a pu, en cas de besoin, remplacer un génitif ambigu. Le sanskrit a l'ablatif *devāt*, avec *-āt* de **-o-od*, distinct du génitif *devasya*, ce qui se comprend fort bien puisque l'élément adjectival ajouté à ce dernier ne saurait, pour le sens, convenir à un ablatif. Mais il a conservé, pour les radicaux en consonne, l'identité première : *vāk* (< **wōk*) « voix » fait, régulièrement, *vācas* (**wōk-és*) aussi bien à l'ablatif qu'au génitif.

Datif et locatif.

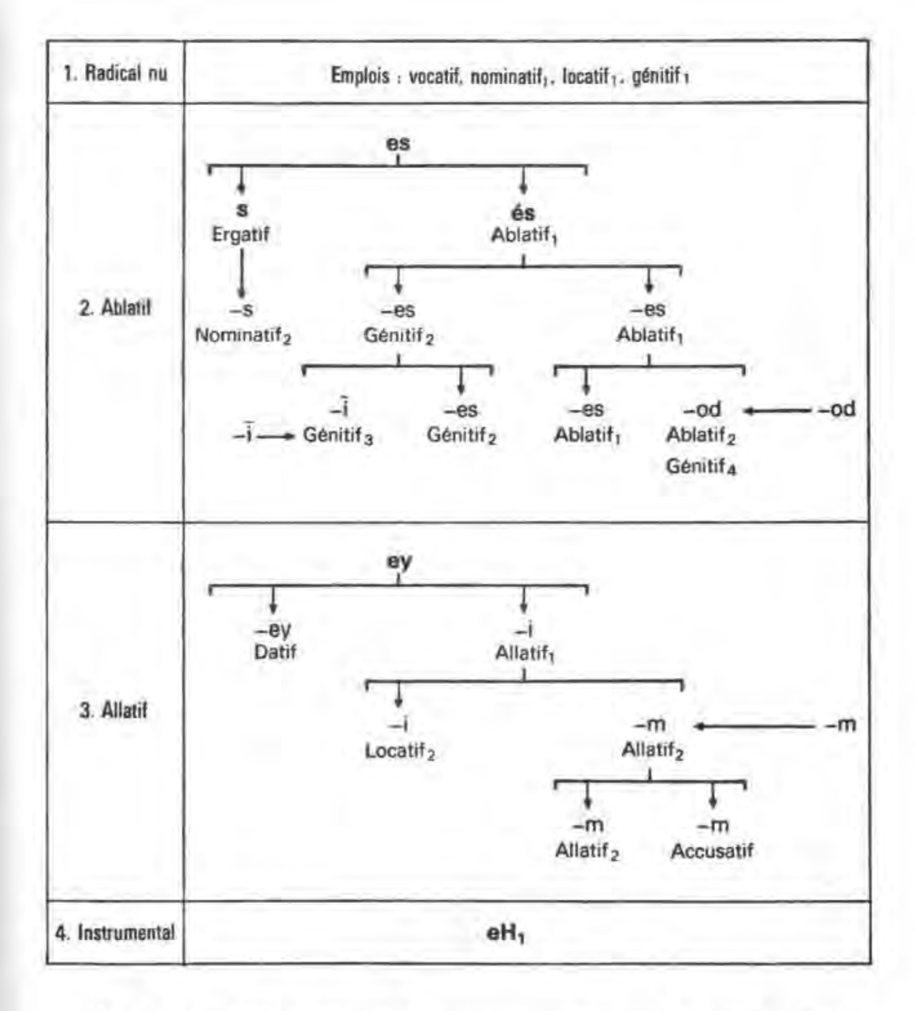
Si l'opposition du nominatif au génitif-ablatif se maintient bien et, le cas échéant, s'accuse, celle de l'allatif au datif, fondée sur l'alternance de *ey* inaccentué à *éy* accentué, est moins ferme. Il faut rappeler d'abord que, sous la pression d'un nouvel allatif en *-m* promis à une belle fortune, l'allatif marqué par *ey* > **-i* suit la voie de la plupart de ses congénères et devient, comme nous l'avons signalé, un locatif n'indiquant plus le lieu où l'on va, mais celui où l'on est. En même temps, se fait rare le locatif initial qui s'identifiait au radical nu (en français *Place de l'Opéra* pour *sur la Place de l'Opéra*).

En sanskrit, l'opposition datif-locatif se maintient bien : on a *vācé* (< **wōkʷ-ei*) au datif, *vācí* (< **wōkʷ-i*) au locatif où l'accent sur la finale est naturellement une innovation. Pour *dēvas*, le datif est *dēvāya*, de **deiwo-oy* plus un *-a* d'origine incertaine ; le *-āy-* y correspond exactement au *-ōi* du datif grec *hippōi* et au *-ō* latin de *dominō* où le *-i* est tombé ; le locatif est *dēvē*, de *deiwo-i*, avec la même finale dans le grec homérique *oíkoi* « à la maison » ; pour les radicaux en consonne, on peut comparer au *vācí* sanskrit le grec *nukt-í* « de nuit ». Mais le grec ne distingue pas le locatif du datif et, comme en latin où l'emploi du locatif est assez limité, on le marque volontiers au moyen d'une préposition. Là où il est attesté en latin, il est marqué par un *-ī* long issu sans doute de *-o-i* à la finale des radicaux à voyelle, dans *bellī* « à la guerre » (< **dwello-i*) par exemple, et étendu, par analogie, à *rūrī* « à la campagne » (< **rūs-oi* au lieu d'un plus ancien **rūs-i*).

Un nouvel allatif.

L'allatif en *-m* est d'apparition relativement récente. Il ne participe pas à l'alternance accentuelle qui nous a valu les oppositions *es* ~ *és*, *ey* ~ *éy*. Ou bien il s'ajoute à une voyelle finale de racine, comme dans sanskrit *agni-m* « feu », *dēva-m* « dieu », lat. *igne-m* (< **egni-m*), *dominu-m* (< **domeno-m*), ou bien, après consonne, il prend une valeur syllabique, **-m̥*, qui aboutit à *-em* en latin : *uōc-em* (< **wōk-m̥*) « voix », et à *-a* en grec : *núkt-a* « nuit ».

L'emploi allatif est encore largement attesté en sanskrit et en grec. En latin il est d'emploi limité, au même titre que le locatif : on dit *Rōmam* « vers Rome », *domu-m* « (aller) à la maison ». On trouve un allatif dans l'anglais *home* (< **koimo-m*) qui s'oppose au locatif *at home*. Dans des circonstances sur lesquelles on reviendra longuement ci-dessous, l'allatif s'est employé pour marquer l'objet du verbe et a pris la dénomination d' « accusatif », due à une traduction erronée du grec *aitiatikē*, cas de l'objet. Cette extension rappelle l'emploi de la



Nota bene : les flèches indiquent des apparentements formels. L'absence de flèche verticale va de pair avec une flèche horizontale qui marque le remplacement d'une forme traditionnelle par une innovation. Les chiffres en indice marquent des stades successifs, mais l'apparition d'un génitif$_3$, par exemple, n'implique pas la disparition du génitif$_2$. L'emploi noté nominatif$_1$ est un véritable nominatif qui sert spécifiquement à nommer, hors syntaxe. Nominatif$_2$ connaît les emplois de nominatif$_1$, mais il est essentiellement le cas sujet.

préposition allative *a* de l'espagnol (opposée à la locative *en*) pour indiquer l'objet là où il pourrait y avoir ambiguïté : *Pilar quiere a Juan* « Pilar aime Jean ».

L'instrumental.

Plusieurs langues indo-européennes possèdent un cas instrumental distinct qui correspond à peu près aux relations couvertes, en français, au moyen de la préposition *avec*, dans *avec sa femme*, *avec un marteau*, *avec calme*. On peut poser au départ une particule eH_1 qui aboutit à une voyelle longue *-ē*, *-ō*, *-ī* ou *-ū*. Il peut y avoir, comme en latin, confusion avec l'ablatif. C'est peut-être ce qui a déterminé en slave et en balte l'utilisation de l'adverbe $*^{m}b^{h}i$, sous sa forme *-mi* ou **-mo* pour distinguer l'instrumental du génitif-ablatif ; l'équivalent grec *phi* est utilisé avec cette même valeur.

On peut résumer le système des cas tel qu'il se manifeste, lorsqu'ils n'apparaissent pas en combinaison avec les nombres pluriel et duel, de la façon du tableau page précédente.

Le pluriel.

Il est très vraisemblable que le pluriel est, en indo-européen, une création relativement récente que l'examen des diverses langues révèle comme en voie d'établissement. En hittite, les seuls cas attestés y sont le nominatif et l'accusatif. Pour les autres cas, on rencontre des formes de singulier là où l'on attendrait le pluriel. Ailleurs qu'en hittite, on trouve des formes qui se correspondent de langue à langue, mais qui distinguent imparfaitement les cas les uns des autres. En latin, par exemple, la désinence *-bus* ($*\text{-}b^{h}os$) recouvre le datif et l'ablatif, cas qui sont, au départ, aussi différents qu'il est possible, puisqu'ils désignent l'un l'approche et l'autre l'éloignement. On trouve à tous les cas un *s* qui paraît bien être la marque du pluriel, mais il n'apparaît pas partout dans les mêmes conditions : final ici, suivi là de la marque du cas.

En conformité avec le sanskrit et le grec, on pose **-es* sans alternance vocalique, donc avec *-e-* et jamais *-o-*[4], comme désinence du nominatif pluriel. Le *-e-* apparaît tel quel après consonne ou sonante finale de radical, mais il se contracte en une voyelle longue avec un *-e-* final de radical. Le *-es* de pluriel ne porte pas l'accent (skrt *vā́cas*) et ne semble pas avoir abouti à la forme tonale qui résulterait de la contraction d'un *-e-* avec un *-é-* accentué suivant (accent dit circonflexe).

4. Sauf, cependant, en germanique où l'on trouve fréquemment des alternances analogiques entre les produits de *e* ~ *o* ; cf. ci-dessus, le nom des Wendes, vieil-angl. *Winedas*, chap. VI, p. 112.

A l'accusatif, l'*s* du pluriel semble avoir été ajouté à la forme de l'allatif-accusatif singulier en *-m*, mais le groupe *-m-s* a évolué en *-ns* avec, le plus souvent, chute du *-n-* et allongement de la voyelle précédente. Là où la finale était *-m̥-s*, le produit régulier est en grec et en sanskrit *-as* avec voyelle brève. L'évolution apparaît ici très différente de celles qu'on a postulées pour les successions *-m-s* ou *-n-s* à la finale des nominatifs singulier. En fait, il y a eu, de part et d'autre, une grande variété d'évolutions proprement phoniques, selon les langues et selon les époques et, ici et là, toute une gamme de restitutions analogiques : là, par exemple, où, dans *-ns*, *-s* devait normalement disparaître, il se voyait automatiquement restitué par des générations suivantes désireuses de marquer les pluriels.

Datif, ablatif et instrumental sont marqués par une forme dérivée de *$^{m}b^{h}i$ avec l'*s* final ajouté. Le sens de cet adverbe, « près de », « à côté de », « autour » semble avoir ici, en quelque sorte, neutralisé les valeurs précises de l'ablatif (« de ») et du datif (« à »). Qu'on se rappelle les valeurs variées de *by* en anglais dans *stand by* « se tenir près de », *done by him* « fait par lui », *by now* « en ce moment », *by boat* « en bateau », *by the year* « à l'année », etc.

Au locatif pluriel, on a dû ajouter le *s* à la forme nue du radical lorsque celle-ci était encore la plus normale pour cette valeur. Plus tard, on a jugé bon d'ajouter à cet *-s-* pluriel le *-i* de locatif. Mais à côté du *-i* du grec *-si*, on trouve le *-su* du sanskrit et le *-ch* (< **-su*) du slave.

Reste le génitif pluriel pour lequel toutes les langues s'accordent pour employer l'équivalent de **-om* après consonne, de *-ōm* pour les radicaux en voyelle, c'est-à-dire *o* + *om*. Le *s* de pluriel, attesté dans le démonstratif avant la désinence en *m*, peut apparaître dans les paradigmes substantifs influencés par ce dernier : c'est le cas des génitifs pluriels latins *dominōrum*, *rosārum* où le *-r-* remonte à *-s-*. Mais, de toute évidence, on a dans *-*om*, *-*ōm*, une forme qui n'a au départ rien à voir avec le pluriel. Il s'agit d'un adjectif en *-o-* figé sous la forme d'un accord avec un substantif neutre. Soit **owi-*, « le mouton » ; l'adjectif dérivé en *-o*, **owi-o-*, aura la valeur d' « ovin », de « relatif aux moutons » ; comme déterminant du neutre *genos-* « la gent, la race », il aura la forme **owi-o-m*, donc **owi-o-m genos* « la gent ovine », « la race des moutons ». Or **owi-o-m* correspond exactement au latin *ouium*, génitif pluriel d'*ouis* « mouton ». L'objection que l'adjectif aurait pu être retenu avec cette valeur sous une autre forme, **owi-o-s*, nominatif non neutre, ou **owi-o-oi*, datif, ne tient pas, parce que tout cas trop marqué comme tel ne pouvait guère prendre la valeur d'un autre cas et que la forme en *-o-m* correspondait, au neutre, au nominatif et à l'accusatif. Elle était, de ce fait, la moins marquée des formes casuelles et probablement la plus fréquente, puisque également celle des accusatifs non neutres.

Le génitif pluriel est une création relativement récente, développée après le départ des Anatoliens et, de ce fait, non attestée en hittite. Elle a toutefois des chances de s'être imposée avant la généralisation de l'accord féminin qui aurait réduit la fréquence de **owi-o-m* au bénéfice du féminin **owi-ā-m*. L'extension de **-om* à l'ensemble des substantifs, quel qu'ait été le genre ou la forme du radical, plaide également en faveur d'une diffusion récente.

Les neutres ont, en matière de pluriel, un comportement particulier. Il semblerait que les neutres, désignant à l'origine des entités passives par nature, aient été peu susceptibles de pluralisation. L'équivalent du pluriel, pour eux, a été acquis par l'emploi d'un collectif en *-eH₂* qui donne *-ā* ou *-a* dans les différentes langues. En grec, ce collectif continue à être traité comme tel, et le verbe dont il est le sujet reste au singulier : *ta zôia trékhei*, mot à mot « les animaux (ça) court » ou, si l'on veut, « le bétail court ». Ce collectif est traité comme un neutre, c'est-à-dire qu'il a la même forme comme sujet ou comme objet, au nominatif et à l'accusatif. Il n'existe d'ailleurs qu'à ces deux cas. Aux autres cas du pluriel, on retrouve le radical du singulier et les flexions ordinaires du pluriel.

En résumé, toutes ces formes de pluriel apparaissent comme le résultat de « bricolages » successifs à partir de formes d'emploi facultatif limitées à l'origine aux nominatifs et accusatifs, voire au simple nominatif avec sa forme caractéristique à voyelle pleine *-es*. Ce qui, à première vue, paraît étrange est que le phonème *s* qui, avec ou sans voyelle précédente, joue, nous l'avons vu, un rôle considérable dans la flexion du nom pour marquer l'agent, l'origine et l'appartenance, réapparaisse ici comme la marque de tout autre chose.

On peut être tenté de poser au départ deux phonèmes sifflants différents dont l'existence a été envisagée ci-dessus. Lorsque ces deux phonèmes auraient menacé de se confondre, les locuteurs auraient, là où les confusions menaçaient, remplacé la sifflante du pluriel par le **-oi*, de même sens, du démonstratif. C'est ce qu'on constate en grec et en latin pour les radicaux en *-o* et en *-eH₂* > *ā*, *a*, représentés dans les démonstratifs masculin et féminin respectivement : le latin a *dominī* « maîtres » et *rosae* « roses » avec *-ī* et *-ae* (de *-oi* et de *-a-oi*) au lieu des *-ōs* et *-ās* attestés dans les autres langues italiques. Toutefois les homonymies casuelles ne semblent pas avoir été très dangereuses, puisqu'on constate que le latin classique s'accommode parfaitement de l'identité formelle totale de *dominī* et *rosae*, l'un et l'autre aussi bien nominatif pluriel que génitif singulier. En anglais contemporain il y a, en dépit de la graphie, identité formelle des formes régulières du génitif singulier, du nominatif pluriel et du génitif pluriel : *the lord's*, *the lords*, *the lords'*. On n'oubliera pas que les nominatifs pluriels

étaient régulièrement accompagnés d'un accord du verbe qui pouvait souvent contribuer à les identifier comme tels.

On peut envisager que le *s* du pluriel ait été, non seulement le même phonème que celui du nominatif-génitif-ablatif, mais qu'ils soient, l'un et l'autre, issus de la même particule *es*. Ceci voudrait dire que la forme accentuée de cette particule a été utilisée non seulement pour le génitif-ablatif, mais, sporadiquement, pour indiquer l'agent au lieu de la forme inaccentuée aboutissant au simple *-s*. Cette forme, constamment accentuée, contrairement au génitif-ablatif qui présente *-os* (grec *podós*) à côté de *-es* (lat. *pedis*), aurait fidèlement conservé le vocalisme *-e-*. Au cours du temps, dans des circonstances particulières, ce nominatif formellement bien distinct de l'autre, aurait acquis une valeur plurielle confirmée par l'accord verbal et aurait été étendu par analogie à tous les substantifs. On entrevoit pourquoi le besoin de distinguer le pluriel du singulier a pu apparaître en priorité au nominatif, la pluralité des éléments actifs s'affirmant sans doute mieux que celle des éléments passifs ou des diverses circonstances. On pourrait penser que la nécessité d'un accord avec une forme verbale de troisième personne du pluriel aurait pu jouer un rôle. Mais il n'est pas certain qu'une telle forme ait existé avant celle du nominatif pluriel. On doit simplement retenir que l'une et l'autre vont de pair.

Il peut être intéressant de relever qu'en basque, le *k* qui caractérise le cas agent (*-k*) et le génitif (*-ko*), est également la marque du pluriel.

L'extension graduelle de la marque de pluriel aux autres cas est documentée et son processus s'explique bien : on comprend que l'objet représenté par l'accusatif et correspondant à un ou plusieurs participants, ait été plus vite atteint que les diverses circonstances qui correspondent aux autres cas. Ce qui pose un problème ici, est la forme sans voyelle du *-s*, parfaitement documentée à l'accusatif et au locatif. Doit-on supposer que l'élément *-e-* du complexe *-es* a été senti comme la marque du cas nominatif, alors que *-s* acquérait la valeur « pluriel » ? Dans ces conditions *-e-* a été perçu comme alternant avec le *-m-* d'accusatif, le *-$^{m}b^{h}i$- d'instrumental, etc., et l'*-s* pluralisant a pu s'étendre aux autres cas.

Le duel.

A côté du pluriel existait, dans certaines des langues indo-européennes anciennes, un duel, bien attesté en sanskrit et en grec, mais également en balte et en slave. Ici, comme pour le pluriel, on a l'impression que la plupart des formes résultent d'extensions analogiques qui n'ont pas été les mêmes de langue à langue. Ce n'est guère qu'au nominatif-accusatif qu'on peut essayer de restituer des formes anciennes, qui sont d'ailleurs celles qu'on rencontre dans le mot pour

« deux ». Le phonème caractéristique paraît avoir été H_3. Ajouté à un radical à finale vocalique, on a *-*e*H_3, d'où -*ō*. Suivi d'une voyelle, on attendrait -*w*-, mais l'analogie peut faire aboutir le produit à *-*ōw*. Le grec a -*ō*; le sanskrit -*āu* < *-*ōw*. Le mot *huit*, qui est le duel d'un terme de mesure qui désigne les doigts de la main moins le pouce, est *oktṓ* en grec, *aṣṭāu* en sanskrit. Après un *i* ou un *u* final de radical, H_3 a pour effet d'allonger cette voyelle, d'où, par exemple, le sanskrit *agnū* « (deux) feux », *šatrū* « (deux) ennemis ». Ces formes sont étendues par analogie aux radicaux en consonnes. Pour ces derniers, le grec présente une désinence -*e* qu'on voit mal comment rattacher aux précédentes.

Les démonstratifs.

On a présenté ci-dessus le démonstratif *so, sā, tod,* au départ strict déictique hors-syntaxe, qui, au cours du temps, s'est procuré toutes les finales susceptibles de marquer ses rapports avec les autres éléments de l'énoncé, et ceci sans qu'il y ait toujours eu identité entre ces finales et celles qui caractérisaient les substantifs, celles que l'on trouve sur le tableau de la p. 195. Ce sont le sanskrit, le grec et le germanique qui ont le mieux conservé les formes anciennes de ce démonstratif et, notamment, l'alternance des consonnes : en sanskrit *sas, sā, tat,* en grec *ho, hē, to,* en gotique *sa, sō, þat-a*. Les correspondances sont là exactes, si ce n'est que le sanskrit présente au masculin un -*s* analogique de nominatif et que le gotique a ajouté une particule -*a* au neutre. Les formes de l'accusatif, respectivement skrt. *tam, tām* et *tat*, grec *tón, tḗn* et *tó*, got. *þan-a, þō, þat-a,* ne posent aucun problème supplémentaire. Le nominatif pluriel masculin **toi* qu'on reconstruit à partir de skrt *tē* et got. *þai* est, nous l'avons vu, à la source des pluriels en -*oi* des radicaux substantivaux en voyelle du grec et du latin. En grec, le *h*- (< *s*-) du nominatif singulier a supplanté le *t*- de **toi*, d'où *hoi*; analogiquement, le féminin correspondant est *hai*. Les autres formes des trois langues sont, surtout en grec, celles qu'on attend de radicaux en voyelle, ou encore des formes diversement renforcées. Au génitif, par exemple, à la forme *te-s/to-s* de base, a été ajouté un suffixe adjectival, -*o*- ou -*yō*-. En germanique, cette forme de base, en gotique *þis*, qui vaut pour le génitif masculin, a été traitée comme un radical auquel on a ajouté des désinences de génitif féminin (*þiz-ōs*), de datif féminin (*þiz-ai*) et de génitif pluriel (*þiz-ē, þiz-ō*). Ailleurs les marques casuelles ont été ajoutées à un radical formé en ajoutant, à **to*, un élément -*sm*- qui est sans doute un ancien **som*- (anglais *same*) avec chute de la voyelle. C'est ce qu'on trouve dans le datif sanskrit *tasmāi* et le got. *þamma* qui valait donc, au départ, « à celui-là-même ».

C'est de ces formes que dérivent celles de la déclinaison dite « forte » des adjectifs allemands. Le *-s* du neutre *gutes* provient régulièrement du *-t-* de *þat-a* (< **tod-*), le *-n* de l'accusatif *guten* du *-n-* de *þan-a* (< **tom-*), le *-m* du datif masculin et neutre *gutem* du *-mm-* de *þamma* (< **to-sm-*), le *-r* des génitif et datif féminin et du génitif pluriel *guter* régulièrement du *-z-* de *þiz-ōs, þiz-ai, þiz-ē* et *þiz-ō*.

C'est de ces formes de **so-sā-tod* qu'on dérive, plus ou moins directement, la flexion des différents démonstratifs dans toutes les langues indo-européennes. Le latin, par exemple, a un système original de démonstratifs où l'on retrouve l'ancien déictique, sous les formes *-te, -ta, -tud*, dans le composé *iste, ista, istud* et où son influence se fait sentir dans l'absence de *-s* de nominatif à la finale de *iste* « celui-ci », *ille* « celui-là », et dans le *-ud* du neutre *illud*.

Les personnels.

Les personnels forment un système bien à part avec, fréquemment, des formes différentes pour la même personne, des désinences de forme inattendue, des indistinctions particulières. Tous ces traits sont assez bien représentés dans les langues contemporaines : le français offre, à la première personne, *je, me* et *moi*, l'anglais *I* et *me*, le russe *ja, mne, menja*, etc. ; l'allemand oppose un accusatif *mich* à un datif *mir*, parallèles au *dich-dir* de la deuxième personne, mais non attestés ailleurs ; à la première et à la deuxième personnes du pluriel, le russe, avec *nas* et *vas*, ne distingue pas le génitif de l'accusatif et du locatif, ce qui rappelle la polyvalence des équivalents latins *nōs* et *uōs*.

Le fonctionnement et l'évolution du système des personnels et ceux des autres éléments nominaux n'ont jamais été tout à fait parallèles. Il y a à cela plusieurs raisons. Une première est la fréquence des emplois hors-syntaxe des différents personnels : *moi ?, moi, j'irais me mêler de ça ? toi, fais attention ! toi, on te le donnera, nous, on veut bien*, en général avec mise en valeur accentuelle. Une autre est que, dans les langues indo-européennes, les personnels sujets sont normalement amalgamés aux verbes. Quant aux autres, ils peuvent être soit enclitiques, comme dans *Jacques me les a donnés*, soit accentuables, ce qui se manifeste encore en français par l'alternance de *me* et de *moi*. Sont également particuliers, les rapports des personnels avec le nombre : le duel, par exemple, peut s'y conserver, voire s'y étendre, là où il a disparu pour les autres nominaux. Le gotique, par exemple, a un jeu complet de personnels duels qui se maintient longtemps dans les anciennes langues germaniques, avec des créations récentes, comme *wit*, évidemment de *wi* + *t*[*wai*] « nous deux ». On trouve tout cela en baltique également, avec, pour « nous deux », le lit. *mù-du* où *-du* représente le numéral *deux*, comme *-t* en gotique.

« Moi et toi » ~ « moi et lui ».

On sait d'autre part que ce qu'on appelle la première personne du pluriel couvre des choses assez différentes. « Nous » peut vouloir dire « moi et toi », mais aussi « moi et lui », sans préjudice de « moi et vous », « moi et eux », et, bien entendu, « moi, vous et eux ». On peut dire que *nous* couvre en fait toutes les personnes, un peu comme *on* qui le fait de façon moins explicite, et ceci explique bien pourquoi, en français contemporain familier, *on* s'emploie pour *nous* sujet, sauf hors-syntaxe : *nous, on s'en fiche !*

La distinction est fréquente, dans les langues les plus diverses, entre « moi et toi » et « moi et lui ». Les formes indiquent souvent que, dans l'équivalent de « moi et toi », la politesse a triomphé et que « toi » est mis en valeur, ce qui en fait une variante du pronom de deuxième personne. Il semble qu'on n'ait pas relevé d'exemples de cette distinction dans des langues indo-européennes, mais on pourrait peut-être en trouver des traces dans les hésitations entre *m-* et *w-* à l'initiale de « nous » dans sa forme accentuée, souvent, au départ, hors-syntaxe. Le sanskrit a ici *w-*, noté *v-*, dans *vayam ;* de même le hittite dans *wes ;* le balte et le slave ont *m-*, avec vieux-sl. *my* et lit. *mes*, mais *w-* en slave au duel *vē ;* le germanique a *w-* dans got. *weis*, all. *wir*, mais on peut s'interroger sur le *m-* de l'allemand dialectal *mir*. Il se pourrait que les formes en *w-*, qui présentent la même initiale que les *deuxièmes* personnes, lat. *uōs*, slave *vy*, sanskrit enclitique *vas*, remontent à un ancien « toi et moi », alors que les formes en *m-* représenteraient un ancien « moi et lui », avec l'initiale *m-* de la première personne du singulier. On pourrait penser que **we-s* (> skrt *vas*) représente *twe* deuxième personne du singulier, avec chute du *t-* initial explicable dans un enclitique exposé à des contextes divers. On peut résumer tout cela dans le tableau suivant :

m…	« moi »	*tw…*	« toi »	
↓				
m…s	« moi et lui » « moi et eux »	*(t)w…s*	« toi et toi » « toi et lui » « toi et eux »	→ sl. *vy* « vous »
↓			« toi et moi »	→ got. *weis* « nous »
lit. *mes*	« nous »			

Restent en dehors les formes atones de 1re personne du pluriel représentées par lat. *nos*, russe *nas*, all. *uns*, v. angl. *us* (< **n̥s-*), hittite *anzas*, et les formes en *ju-* de la 2e personne du pluriel du sanskrit, du germanique et du balte.

Vont de pair avec les personnels, le réfléchi qui, naturellement,

puisqu'il suppose le rappel, dans une fonction non sujet, d'un nominatif présenté antérieurement, n'a pas de nominatif. La base est *swe-* tout à fait parallèle au *twe-* de deuxième personne. En latin, par exemple, les formes *sē, sibi* riment avec *tē, tibi*.

Les possessifs.

Les possessifs ne peuvent être dissociés des personnels. Ils apparaissent souvent comme des génitifs de ces personnels ou, ce qui en est l'équivalent, comme des adjectifs dont ils dérivent : « mon fils » vaut évidemment « le fils de moi ». La forme initiale des uns et des autres est certainement pour 1. et 2. **me/mo* et **t(w)e/t(w)o*, particulièrement bien attestés en celtique. Dans une des rares phrases qui nous restent du gaulois, on a, de la bouche de la mère de Saint Irénée, pour le confirmer dans son martyre, le vocatif *mo gnāte* « mon fils ». En v. irlandais on a *mo* et *do* (forme intervocalique de **to*). Le hittite a *me-s*, *te-s* avec un *-s* analogique de nominatif. Nous retrouverons, ci-dessous, ces formes lorsqu'on traitera des désinences verbales.

Pour le pluriel, on trouve essentiellement divers dérivés, souvent avec un suffixe *-r-* qu'on a, par exemple, dans got. *unsar* (< **n̥s-ro*), all. *unser*, angl. *our* (< **ūsr-*). Cet *-r-* peut être ce qui reste de *-tro-* ou *-tero-* qu'on a dans le latin *nostrī, nostrum* (< **nos-tr-*, plus la marque du génitif singulier *-ī* ou pluriel *-um*) et le grec *(he)métero-*. Ce suffixe est bien connu ailleurs, par exemple dans les désignations de la droite et de la gauche et dans le mot *autre*, latin *alter*, got. *an-þar*, all. *ander* où il marque l'altérité. On a sans doute, au départ, un simple personnel, « nous », c'est-à-dire « [moi et] autre que moi », comme l'indiquent les marques additionnelles de génitif en latin. Toutefois, à côté des dérivés, on trouve, en hittite, des enclitiques *-me-s*, **-te-s*(?), *se-s* avec ce qui peut être ici un *-s* de pluralisation. Les deux premiers se retrouvent dans les désinences verbales du latin *-mus* (**-mos*) et *-tis* (**-tes*) respectivement.

On peut donc poser, à date ancienne, des formes *me/mo, te/to, -me-s/-mo-s, -te-s/-to-s*, qui valaient pour *me* ou *mon*, *te* ou *ton*, *nous* ou *notre*, *vous* ou *votre*, mais nullement pour les formes hors-syntaxe de mise en valeur. Accolées à un substantif, elles prenaient une valeur possessive, rappelant ce qui se passe en composition ; le gaulois *mo gnāte* « fils de moi » étant parallèle à grec *Dēmo-sthénēs* « Démosthènes », littéralement « force du peuple ».

Les adjectifs.

Ils ne se distinguent pas au départ des substantifs. Les dérivés en *-o* et *-yo* ont très fréquemment des valeurs et des emplois adjectivaux,

mais, anciennement, cela n'implique pas deux classes différentes dont on ne passe de l'une à l'autre que par une opération de transfert. En grec, le « passif » *tómos* se traduira par « portion », « morceau coupé » et l' « actif » *tomós* par « tranchant », l'un et l'autre étant des radicaux en voyelle : *tomo-*. Ceci, qui paraît assez naturel aux usagers contemporains, démarque nettement l'indo-européen de beaucoup d'autres langues où nos adjectifs se traduisent par des verbes : il n'y a pas de « grand », mais uniquement des « être grand », nos épithètes (« grand homme ») se présentant comme des relatifs (« homme qui est grand »).

Les cardinaux.

Nous les avons déjà rencontrés à maintes reprises. On signalera simplement ici qu'ils ne se fléchissent dans les langues attestées que jusqu'à quatre. Au-delà, ils sont invariables. Dans les langues contemporaines les plus évoluées, l'invariabilité est totale, sauf pour *un* où seul l'anglais, qui a éliminé les genres, a pu aller jusqu'au bout, avec *one*.

Pour « un » on a dû hésiter entre **oi-*, plus un suffixe variable, *-ko-* en sanskrit, *-wo-* en iranien et en grec, *-no-* à l'ouest (d'où, régulièrement, lat. *ūnu-s*, all. *ein*), et *sem-* « unique », « soi-même », qui est à l'origine du grec *heî-s* (< **sem-s*), féminin *mía* (**sm-iya*), neutre *hén* (**sem*) et qui est abondamment représenté ailleurs, par exemple dans le lat. *simplex*, littéralement « plié une fois », de la liste des *duplex* « double », *triplex* « triple », etc.

De « deux » et « trois » on peut dire qu'ils rimaient respectivement avec les duels et les pluriels : skrt. *du-āu* comme *dēv-āu*, *tray-as* comme *agnay-as*. Il faut citer le radical **dwi-* pour « deux » attesté dans le féminin lituanien *dvi* et dans le skrt. *dvi-s*, le latin *bi-s* et le grec *dí-s* « deux fois ». On peut penser à un ancien *dwiH-* avec perte de *-H-* en position prévocalique, qui correspondrait aux duels sanskrits en voyelle longue.

« Quatre » semble avoir été formé à partir de $*k^{w}et$- auquel s'est ajouté un élément *-wōr* qui alterne, au féminin, avec un élément *-sōr*, qu'on connaît bien dans **swe-sōr* « sœur », la femme qui appartient à la grande famille, marquée par *swe* « soi-même », donc « la femme de chez nous », et dans **uk-sōr*, l'épouse, la femme qui vient d'ailleurs, sans qu'on puisse exactement identifier *uk-*, dont on ne trouve trace qu'en latin et en arménien.

« Cinq », reconstruit comme $*penk^{w}e$, ne saurait être dissocié d'angl. *finger*, all. *Finger*, *fing-* correspondant régulièrement à $*penk^{w}$- devant consonne. Le latin et le celtique ont régulièrement changé $p...k^{w}$ en $k^{w}...k^{w}$, ce qui explique le latin *quinque* (dans *quinconce*) et

l'irlandais *cóic ;* fr. *cinq* et ital. *cinque* supposent une dissimilation de $k^w ink^w e$, en $kink^w e$, c'est-à-dire l'inverse de ce qui avait eu lieu précédemment. Le grec *pénte* (d'où *pentagone*) est tout à fait régulier.

L'initiale de *six* pourrait avoir été *ksw-*, mais la plupart des langues ont la forme **seks* suscitée ou favorisée peut-être par le *s-* du cardinal suivant **septm̥*. Dans ces formes, qu'à partir de « cinq » on apprend en comptant, il y a souvent analogie par anticipation : *vingt(e)-cinq*, comme *trente-cinq*.

Nous avons déjà rencontré *huit*. Quant à *neuf*, on pourrait y voir une forme de **new-* « nouveau » avec le *-m̥* final sur l'analogie du suivant **dekm̥* « dix », d'où **newm̥* (> lat. *nouem*). Le français aurait, dans ce cas, retrouvé l'identité formelle de *neuf* « nouveau » et du cardinal. Cela voudrait dire qu'on avait compté tout d'abord jusqu'à quatre, sur les doigts sans le pouce, quatre objets étant d'ailleurs ce qu'on peut percevoir comme distincts sans l'aide du langage, puis jusqu'à deux fois quatre, donc huit, le nombre suivant commençant une nouvelle série.

Les premiers cardinaux de la liste sont traités comme des adjectifs. Puis, à un certain point, on passe aux substantifs : *cinq hommes*, mais *un million d'hommes*. Mais ce point varie de langue à langue. En français il se place entre *mille* et *million*, en anglais entre *ninety*, 90, et *(one) hundred*, 100, en russe entre *četyre*, 4, et *pjat'*, 5.

LE VERBE

La « conjugaison ».

Conjuguer c'est, étymologiquement, mettre ensemble sous le joug. Pour comprendre ce qu'est le verbe indo-européen, il ne faut pas voir, dans la conjugaison, la mise ensemble de différentes personnes pour constituer un temps, ce que suggéreraient les langues contemporaines, mais le procès par lequel *des mots différents* ont été, à une époque antérieure aux textes les plus anciens, groupés pour en former *un seul*. Sous une forme très limitée, le processus s'est poursuivi jusqu'à une date tardive : le verbe *être*, en français, combine des formes empruntées à des racines distinctes. Certaines, comme celle de *est*, *sont*, c'est-à-dire **es*, et celle de *fut* étaient déjà « conjuguées » en latin, mais celle de *était*, tirée du latin *stare*, était, dans cette langue, un autre mot avec une flexion complète dont le français n'a retenu que l'imparfait et le participe parfait, si l'on fait abstraction du rare *ester*.

Un autre exemple est *aller* qui combine deux verbes latins bien distincts, *eo*, *i-re*, proprement « aller », *uado*, *uadere* « s'avancer », au départ, peut-être, « passer à gué », et une forme un peu obscure qu'on rattache souvent au latin *ambulo*, *ambula-re* « se promener ».

Du latin au français, on ne relève que des traces sporadiques du

phénomène, mais d'autres langues, comme le slave, ont, à date assez récente, renouvelé leur système verbal en combinant un verbe simple, russe *pisat'* « écrire » avec un composé *na-pisat'*, de même sens. Les deux formes fléchies *pišu* et *na-pišu* fonctionnent, la première, comme présent : « je suis en train d'écrire », l'autre comme futur « j'écrirai (et terminerai) ».

Le phénomène de « conjugaison » qui a abouti aux systèmes verbaux qu'on relève dans les langues indo-européennes les plus anciennement attestées est, en principe, le même dans toutes les branches de la famille. Cependant les résultats varient d'une langue à une autre, chaque verbe grec, par exemple, groupant ses formes sous trois rubriques : « présent », aoriste et parfait, alors que le latin n'en connaît que deux : *infectum* et *perfectum*.

Accompli et procès en cours.

Pour comprendre la genèse et l'évolution du système verbal de l'indo-européen, il faut partir d'un état où l'on distinguait entre la constatation d'un état de fait résultant d'une opération antérieure et celle d'un procès en cours. Ce système binaire se retrouve à la base du système verbal de très nombreuses langues un peu partout dans le monde. Il est attesté dans beaucoup de créoles et de langues africaines qui en sont les substrats probables. Quand on examine comment les enfants français acquièrent la maîtrise de leur langue, on s'aperçoit qu'une des premières distinctions acquises est celle qu'on relève entre « papa (est) parti » (il n'est plus là) et « papa s'en va » (il est en train de partir). La première de ces formes a été désignée comme le parfait. Mais, à l'usage, le terme est devenu ambigu car il est souvent conçu comme désignant le passé. C'est pourquoi il vaut mieux, quand il y a danger de confusion, y référer comme à l'accompli : *il a fini*, sans aucune addition, est un présent accompli, *il avait fini* un passé accompli ; mais *il a fini hier à cinq heures* est simplement un passé, et l'identité formelle de *il a fini* dans les deux cas explique le passage du sens de *parfait* à « passé ». Après cette mise en garde, nous conserverons ici le terme traditionnel de **parfait** pour désigner l'accompli, la constatation d'un état de fait résultant d'une opération antérieure.

Comme l'interprétation des formes, qui fonde notre reconstruction en la matière, pose des problèmes délicats dont la solution risque à tout instant de détourner l'attention du but à atteindre, nous allons commencer par esquisser à grands traits l'évolution qu'on suppose, pour l'étayer, dans un second temps, au moyen des formes attestées.

Nom et verbe.

Soit un très ancien état de langue où la distinction entre nom et verbe n'est pas encore établie. Ceci veut dire qu'on ne distingue pas formellement entre *courir* et *course*, entre *craindre* et *crainte*, c'est-à-dire entre des termes qui se réfèrent à une même réalité, mais dans des contextes différents. Comme, en fait, « nom » s'oppose à « verbe », il vaut mieux dire ici qu'il n'y a, dans un tel état de langue, ni noms ni verbes. Mais nous avons intérêt à concevoir ces unités indifférenciées comme des noms, parce qu'on suppose qu'elles présentent un trait qui caractérise, pour nous, les noms par opposition aux verbes, à savoir l'indifférenciation du passif et de l'actif : si nous disons *il craint*, nous spécifions que la crainte est *ressentie* par la personne représentée par *il*, mais, dans *la crainte de l'ennemi*, nous ne précisons pas si la crainte est *ressentie* par l'ennemi ou *inspirée* par lui. Soit encore la paire *tuer* et *meurtre :* avec le verbe *tuer*, je pourrais distinguer entre *il tue*, où *il* renvoie à un individu *actif* et *il est tué*, où *il* renvoie à un individu *passif*. Rien de tel avec *meurtre :* le *meurtre de l'homme* ne permet pas de savoir si l'homme en cause a donné ou subi la mort, a été actif ou passif.

« Noms d'action » et « noms d'auteur ».

Un début de différenciation a pu apparaître lorsqu'à partir d'une forme de base, neutre quant à l'orientation de l'action vers celui qui la subit, le patient, ou celui qui est à sa source, l'agent, on a dérivé (par adjonction d'un *-t*) ce que nous nommons un « nom d'action » permettant de préciser l'identité de l'agent et du patient. L'agent a été marqué comme tel par la particule indiquant l'origine (**es*) sauf dans le cas de personnels où l'on disposait de formes simples, avec la valeur génitivale des radicaux nus, assimilables à nos adjectifs possessifs (**mo*, **to*).

Un autre processus de dérivation (par adjonction de la voyelle *e/o*) a visé simplement à marquer, sans plus, un rapport avec l'entité indiquée par la base. C'est ce que nous serions tenté de désigner comme une dérivation adjectivale. Nous avons eu, ci-dessus, maintes occasions de mentionner l'apparition d'adjectifs en *-o*, comme l'ordinal lat. *decimus* (**dekm̥-o-*) à partir du cardinal *decem* (**dekm̥*). Toutefois une tendance s'esquissera vite à opposer, dans ces formations, l'action, avec l'accent sur la base, et l'acteur, avec l'accent sur le suffixe. Dans ce dernier cas, lorsque la forme en *-é* (final accentué) désigne un être humain, la valeur sera plutôt celle d'**auteur**, celui qui

est ce qu'il est parce qu'il a agi d'une certaine façon dans le passé[5]. **L'acteur** proprement dit sera plutôt formé au moyen d'un suffixe en *-r*, à partir du nom d'action en *-t* envisagé ci-dessus, comme dans lat. *uictor* « celui qui vainc ». La forme du nom d'auteur s'emploiera, dans ce que nous appellerions une « proposition nominale », sans verbe, équivalente à nos propositions à copule : *il est fatigué, il est le père de six enfants, père* étant conçu comme *l'auteur* des jours des six enfants. Si la personne à désigner comme l'auteur est exprimée au moyen d'un pronom, celui-ci prendra naturellement une forme nue, qui coïnciderait avec la forme hors-syntaxe si celle-ci n'était souvent renforcée de particules diverses. Le rapport entre les deux éléments en présence n'est pas de détermination de l'un par l'autre, mais d'identité. On aura en quelque sorte une équation « auteur » = « moi » sans marque de dépendance syntaxique. On attend donc de part et d'autre une forme nue, sans aucune désinence.

Nous avons, à ce point, d'une part, un « nom d'action » en *-t* combinable avec une désignation d'agent marquée par **(e)s* ou un pronominal quasi possessif, d'autre part, un « nom d'auteur » qu'on attribue à une entité représentée par un radical nu. Dans ce dernier cas, on ne s'attend pas à ce qu'un autre élément participe à l'équation sur le même plan que les deux membres : dans « lui- » = « père de six enfants », « enfants » vient préciser « père » et n'affecte pas la structure de l'équation. Il ne s'agit pas du tout de ce qui serait, en français, « le père a engendré six enfants » où « enfants » participe, tout comme « père », mais avec un rôle différent, à la structure syntaxique de base. Dans l'état de langue que nous envisageons ici, « enfants », comme détermination de « père » s'exprimerait au moyen d'un radical nu juxtaposé, le génitif$_1$ de notre tableau, p. 195.

Au contraire, dans les propositions à « noms d'action », il pourra y avoir, outre l'agent marqué positivement comme tel, un participant passif, celui qui subit l'action. Son rapport au « nom d'action » sera le plus direct, celui qu'on trouve, dans les composés des langues indo-européennes attestées, marqué, ici aussi, par la simple juxtaposition. Mais il n'y a plus, dans ce cas, équation, mais rapports de détermination d'un noyau, le « nom d'action », par deux satellites, l'agent positivement marqué par *(e)s*, le patient sous la forme du radical nu.

En résumé, on a, à ce temps de l'évolution, deux types bien distincts :

5. Le terme m'a été suggéré par Emile Benveniste, en 1950, dans une conversation où nous constations l'identité de nos vues relativement à l'origine du parfait indo-européen. Je n'ai pas connaissance d'un exposé écrit de ses vues en la matière.

1) homme = auteur de meurtres ;
2) mise à mort
↑ tigre ↑ par homme

Notons que « tigre » du second à la même forme nue que « homme » dans le premier.

Apparition d'un sujet.

La « conjugaison » va naître de la combinaison en un système verbal de ces deux types de construction, auxquels s'en ajoutera, comme nous le verrons ci-dessous, une autre impliquant des noms d'état à participant unique. Dans l'équation « personne en cause » = « auteur », où les deux termes réfèrent à un seul et même individu, le premier élément va être conçu comme l'agent des actions dont on constate le résultat. Il recevra finalement, de ce fait, la marque *-s*, celle que nous pouvons désigner maintenant comme le nominatif au sens traditionnel du terme de cas sujet (notre nominatif$_2$, ci-dessus, p. 195). Du même coup, le second élément sera interprété comme désignant ces actions. Ce nominatif va désormais être présent dans les deux types fondamentaux d'énoncés :

1. (par) homme → meurtres commis (actions conçues comme accomplies)
2. (par) homme → mise à mort ← tigre (action conçue en elle-même).

Du fait de son emploi obligatoire dans tout énoncé, l'ancienne forme d'agent va prendre le caractère d'un « sujet », ce au sujet de quoi on dit quelque chose. Elle n'aura donc plus à expliciter ses rapports syntaxiques, parce qu'on attend que le reste de l'énoncé prenne position par rapport à elle. Il en découle deux conséquences :

1. Cette forme va pouvoir s'employer pour *nommer* quelqu'un hors-syntaxe en remplacement du radical nu, d'où sa désignation traditionnelle comme « nominatif ». On va également être tenté de l'utiliser dans l'appel et elle va, de ce fait, tendre à éliminer la forme dite « vocatif ».

2. L'ancienne particule **(e)s*, devenue la simple désinence *-s*, va être sentie comme superfétatoire et elle ne va plus être automatiquement renouvelée par analogie chaque fois que l'évolution phonétique l'élimine dans certains contextes. La tendance s'est poursuivie à travers des millénaires : elle a dû commencer avec les radicaux en *-r* du type **pətēr*, il y a cinq millénaires ou plus, et n'aboutir définitivement en français qu'au moment où, vers la fin du moyen âge, le cas sujet en *-s* du singulier est abandonné au profit du cas objet sans *-s*, pour ne survivre que dans des isolés, comme *fils*, ou dans la graphie de noms

propres masculins, comme *Georges* et *Jules*. Il n'y a plus guère aujourd'hui que l'islandais (sous la forme d'un nominatif en *-r*) et les langues baltes pour traîner encore cette encombrante marque d'une relation syntaxique disparue.

Il nous faut maintenant illustrer nos dires au moyen de formes attestées ou reconstruites.

Le « nom d'action ».

Considérons d'abord le « nom d'action ». Nous partirons d'une base *g^{hw}*en*. Elle renvoie, au départ, à des violences susceptibles de causer la mort. Elle est largement attestée avec un élément suffixal en *-t-* diversement renforcé au cours du temps. Avec un vocalisme zéro, qui laisse supposer une accentuation du suffixe, donc *g^{hw}*n̥-t́—*, on trouve skrt. *hatis* « mise à mort » et, en germanique, une forme *gund-* qui y désignait la bataille et la guerre avant que le lat. *gregarius* « (simple) soldat » fournisse l'ancêtre de l'all. *Krieger* « guerrier », d'où, par soustraction du suffixe, *Krieg* « guerre ». En sens inverse, le français *guerre* dérive d'un emploi argotique d'un germanique *werra* « la rixe, la bagarre ». *Gund-* est représenté en français dans *gonfalon* ou *gonfanon*, « l'étendard (*-fanon*, all. *Fahne* « drapeau ») de guerre », sans parler de prénoms comme *Gontran*, dialectalement *Gondran*, et *Cunégonde*.

A la base *g^{hw}*en-t* on ajoute *-mo*, d'où le sens de « ma frappe » ou « frappe de moi » avec le même rapport syntaxique que celui qui est signalé ailleurs par la particule **(e)s*[6]. Dans les deux cas, on marque l'origine ou l'acteur. *g^{hw}*en-t-mo* évoluera en *g^{hw}*en-mo* qu'on retrouve, par exemple, dans le vieil-irlandais *gonim* « je blesse », « je tue ».

A la deuxième personne, on a *g^{hw}*en-t-to* où *-t-t-* évolue régulièrement en *-s*[7]. Il en résulte *g^{hw}*en-so* « frappe de toi ».

A la troisième personne, on a la forme nue *g^{hw}*en-t* « (il y a) frappe », sans rappel d'un acteur impliqué par le contexte ou la situation. Lorsque la forme, devenue proprement verbale, aura été intégrée à un paradigme, la voyelle *-o* y sera ajoutée par analogie avec celle des première et deuxième personnes, *-mo* et *-so*.

Les formes de pluriel attestées, la première en *-me/os* (lat. *-mus*), la seconde en *-te/os* (lat. *-tis*), sont évidemment des pluralisations par *-s* des *-mo* et *-to* des singuliers. Dans *-te/os*, le *-t-* a été conservé, contrairement à celui du singulier passé à *-s-*, parce qu'ajouté à une base, désormais conçue comme g^{hw}*en-*, commune aux formes en *-mo*, *-so* et *-to*.

6. On s'inspire directement, ici, du traitement d'André VAILLANT dans l'ergatif indo-européen, *BSL* 36, p. 93-108.

7. Voir, ci-dessus, p. 167-168.

Les formes en *-nt-* du pluriel de troisième personne sont, sans doute, relativement récentes, la pluralité des acteurs ayant initialement été impliquée par le contexte. En français contemporain, *il donne* et *ils donnent* sont de parfaits homophones et, sauf certains Français du nord qui font, au pluriel, sentir le *-t* final, les usagers ne semblent pas éprouver le besoin d'une distinction. Ces formes en *-nt-* sont sans doute à rapprocher de celles du participe de même forme : « ils frappent » = « eux frappant ».

Le « nom d'auteur ».

Soit, maintenant, le « nom d'auteur ». Nous partons ici d'une base **weid-* qui implique « vision ». On en tire un « adjectif » en *-e/o* accentué sur le suffixe, d'où **widé*, mais analogiquement **woidé* sur le modèle d'un *g^{hw}*né* (de *g^{hw}*en*) imprononçable et passé à *g^{hw}*oné*. Avec recul ultérieur de l'accent, c'est le grec *woíde* et, très exactement, l'all. *(er) weiss* qui n'a pas plus de *-t* de troisième personne que les formes du même type *(er) kann*, *(er) mag* dont le *a* du radical correspond précisément au *-o-* de *woide*. Le sens de *woíde* et de *weiss* est « il sait » et « il sait parce qu'il a vu ». En grec *woíde* est le « parfait » de *weídō* « je vois ». A l'origine, c'est « celui qui a vu », c'est-à-dire « le témoin ». Ici encore, il n'y a pas, au départ, de pronom indiquant de qui il s'agit, contexte et situation suffisant parfaitement. Pour les personnes exprimées, la première et la deuxième, on doit supposer que l'élément pronominal a attiré l'accent de telle sorte que le radical se réduit à *woid-*. Pour la première, on peut poser *-H_2é* et, pour la seconde, *-tH_2é*, en grec 1. *woída* et 2. *woîstha* où *-d + t-* est passé régulièrement à *-s-* et où le *-t-* a été rétabli par analogie. Le traitement de *-H_2-* comme *h* et, finalement, comme marque de l'aspiration du *t* est, on l'a vu, sporadique en grec.

Les marques personnelles au pluriel sont, ou bien comme en grec, calquées sur celles des formes verbales dérivées du « nom d'action », ou bien, comme en sanskrit, difficilement identifiables par la comparaison. Il faut faire exception d'une finale en *-r* qui caractérise la 3ᵉ personne du pluriel et que nous retrouverons plus loin.

Ce que nous avons désigné comme le nom d'auteur n'est qu'une des valeurs des formations « adjectivales » de radical en voyelle. Accentuées sur l'initiale, elles valent pour une action accomplie. A partir de *g^{hw}*en*, on a *g^{hw}*ón-o-* « meurtre », devenu régulièrement, en grec, *phóno-s*, distinct de *phonḗ* (*g^{hw}*on-eH_2*) « action de tuer ». De même, le grec *tómos* désigne le résultat de l'action, ce qui est coupé (cf. le français *tome* pour une partie distincte d'un ouvrage), en face de *tomḗ* « action de couper ». Accentué sur le suffixe, la valeur est active, *tomós* c'est ce qui coupe, qui est coupant.

Le parfait, une fois constitué comme un élément de conjugaison, s'est vite révélé, dans les différentes langues, comme insuffisamment distinct de ses partenaires d'autre origine. Les locuteurs ont alors cherché à le différencier supplémentairement par extension analogique de formes épisodiques, c'est-à-dire limitées à certains verbes ou à certaines personnes. Un des traits assez répandus est le redoublement de la première syllabe plus ou moins déformée, comme dans lat. *momordi* « j'ai mordu ». Mais, le plus souvent, on se contente de répéter la consonne initiale en la faisant suivre de *-e-*, comme dans grec *dédorka*, parfait de *dérkomai* « je regarde », en sanskrit *dadárça*. En germanique, où le redoublement est attesté, c'est cependant surtout le vocalisme *-o-* du radical qui a été mis à profit sous sa forme régulière *-a-*. Le parfait y a pris totalement la valeur d'un passé, comme dans le passage du latin aux langues romanes, et il y figure comme prétérit, par exemple dans angl. *he sang* « il chanta ». Quelques parfaits y ont pris la valeur d'un présent, comme *he can*, *he may* et l'archaïque *(he) wot* (< **wait-*) correspondant au gr. *woíde*.

Comme on l'a vu ci-dessus, p. 152, le latin a fait grand usage du suffixe *-u(ī)* dérivé d'un ancien H_3 à la finale de certains radicaux comme **stre*H_3- « étendre », lat. *strāuī*, *$b^h le H_3$ « souffler », lat. *flāuī* ; ce suffixe s'est généralisé dans angl. *strew* et *blow*.

La formation régulière du grec offre *-k(a)* qui doit provenir de formes de première personne de radicaux en -H_2 ou -H_3 où la « laryngale » était en contact avec le H_2 de l'élément pronominal suivant.

Il y a eu, dans différentes langues, création d'un passé du parfait, ce qu'on désigne en général comme un « plus-que-parfait », parce que, du fait d'une évolution qui paraît inéluctable, le parfait finit par s'employer sporadiquement ou généralement avec valeur de passé. C'est le cas du parfait du français contemporain qu'on désigne sous le terme de « passé composé ». On a vu que *il a fini* est un présent accompli, tout comme *j'ai mangé*, mais que *il a fini hier à sept heures*, ou *j'ai mangé à 10 heures*, sont de simples passés qui remplacent aujourd'hui, dans l'usage quotidien, les anciens prétérits *il finit, je mangeai*. Le parfait latin, qui a donné le passé simple d'aujourd'hui, avait la même double valeur : *uīxit* est ou bien « il est dans l'état de quelqu'un qui a cessé de vivre » c'est-à-dire, au présent, « il est mort », ou bien, dans le passé, « il a vécu ».

Les éléments de la « conjugaison ».

Le hittite n'a pas participé à ce qu'il faut comprendre proprement comme la « conjugaison ». Les formes issues du nom d'auteur n'y sont pas devenues des parfaits qui s'adjoignent à d'autres verbes pour

former de nouvelles unités, mais ils ont formé un type particulier de verbes indépendants. Soit un nom d'auteur **sok*^{w}*é*; il est celui qui a suivi (cf. latin *sequor* « je suis ») des yeux, qui a donc vu et qui sait, un peu comme *woíde* en grec. En hittite, il devient un verbe indépendant avec le sens de « savoir ». En combinaison avec la marque *-H_2e* de première personne, plus une marque *-i* de présent, on a une forme **sok*^{w}*H_2ei* qui aboutit à *sakhi*. Le gotique connaît la même forme, devenue régulièrement *sah*^{w}, mais avec la valeur d'un passé, donc sans le *-i*, « j'ai vu », angl. *I saw*.

En germanique, la conjugaison s'est limitée à la mise ensemble du produit du « nom d'action » et de celui du « nom d'auteur », celui-ci ayant finalement fait évoluer sa valeur d'accompli en celle de passé. En latin et en celtique, il y a eu également passage au passé de l'ancien « nom d'auteur », mais certaines formes dérivées du « nom d'action » ont également contribué à la création de ce temps.

Événements et situations.

Dans d'autres langues, comme le grec et le sanskrit, le phénomène de « conjugaison » a consisté à joindre, non pas deux verbes différents pour en faire un seul, mais trois. D'une part, celui qui a donné le parfait, issu du « nom d'auteur » ; d'autre part, deux verbes issus, par dérivation divergente, de l'ancien « nom d'action ». Une fois « conjugué », un de ces verbes donne ce qu'on appelle traditionnellement le « présent », l'autre ce qu'on désigne comme l' « aoriste ». Le terme d'aoriste, qui n'évoque pas grand-chose de précis pour le lecteur contemporain, n'est pas dangereux. « Présent », au contraire, suggère immédiatement un temps et l'on s'étonne un peu d'apprendre qu'il existe un passé du « présent ». Pour éviter les confusions, on a sans doute intérêt, ici, à employer des désignations moins ambiguës et plus descriptives des faits en cause, même au prix d'un alourdissement du vocabulaire. Il faut comprendre que les deux verbes, dérivés du « nom d'action », ont été l'un, la référence à un *événement* ponctuel ou envisagé comme tel, l'autre à une *situation* dynamique dans laquelle se déroule un procès. Lorsqu'un verbe, quelle que soit son origine, désigne, non une action, mais un état, il se rangera parmi les situationnels et, vu son sens, il pourra être difficile de lui trouver un partenaire événementiel. C'est ce qui se manifeste en grec dans le cas du verbe d'état par excellence, le verbe *être*, qui n'a pas d'équivalent événementiel. En latin et ailleurs, on a rempli la « case vide » au moyen d'un verbe du sens de « devenir », d'où en français, *il fut* en face de *il est*. A y réfléchir, on constate que *il fut tué* ne fait pas grand sens si l'on prête à *fut* la valeur existentielle qui est celle d'*être*. L'équivalent allemand *er wurde getötet*, avec un auxiliaire du sens de

« devenir », se comprend, mot à mot, comme « il devint tué ». En anglais, *he got killed* fait précisément référence à un événement, alors que *he was killed* suggère aussi bien l'action accomplie.

Un événement, c'est-à-dire l'aboutissement d'un processus, et non ce processus lui-même, n'est guère susceptible, en tant que tel, de présenter diverses modalités. C'est pourquoi, lorsque s'est produite la conjonction de trois verbes en un seul, il n'y a pas eu le choix entre beaucoup de mots différents pour venir remplir la case de l' « aoriste » événementiel. En grec, par exemple, on y trouve — outre quelques radicaux sous leur forme la plus simple, comme *pio-* pour « boire » avec la valeur qu'il a dans « il but » — essentiellement des formes en *-s-* au moyen desquelles on produira, à chaque instant, de nouveaux « aoristes » comme *é-lu-s-a* « j'ai dételé ». En latin et en celtique, langues qui n'ont pas concrétisé l'opposition de l'événementiel et du situationnel, cette forme en *-s-* a été annexée au parfait : lat. *uexit* ($<$ **weg^h-s-H$_2$e-i + t*) « il a transporté » ; on trouve, en gaulois, une forme *legasit*.

Il en va tout autrement d'une situation qui évolue. Sans cesser d'être évolutive, elle peut impliquer un point de départ ou un point d'arrivée : à côté de *porter*, on a *emporter* et *apporter*. Alors qu'on peut symboliser l'événement par un point sans flèche, la situation évolutive pourra combiner point et flèche. Donc



Il y avait, avant que ne se produise la « conjugaison », des façons très diverses de dériver un verbe nouveau à partir d'une racine ou d'un radical existant. On pouvait utiliser un suffixe comme *-yo-*, comme *-eH$_1$-* ou *-eH$_2$-*, comme *-sko-* ou *-esko-*, ou encore un *-n-*, soit infixé (lat. *li-n-quo* « je laisse » en face du parfait *liqui*), soit suffixé et diversement élargi (grec *deík-n-u-mi* « je montre », *dú-n-ā-mai* « je peux ») ou répété (grec *la-n-th-án-ō* « je suis caché » ; cf. sans *-n-* l'équivalent latin *lateo*), ou encore le redoublement de la consonne initiale (grec *dí-dō-mi* « je donne »). Les valeurs de ces divers procédés ont pu varier d'une langue à l'autre, d'une époque à la suivante, mais tous semblaient impliquer ce que nous avons désigné comme une situation.

Au cours du processus d'appariement, que suppose la « conjugaison » en grec ou en sanskrit, d'un événementiel avec un des divers situationnels, tel événementiel a pu s'accorder avec tel situationnel dans une langue, tel autre dans l'autre langue. Pour la notion de « laisser », l' « aoriste » événementiel est le même en grec et en sanskrit, à savoir **e-lik^{w}o-* qui apparaît comme *élipo-* d'un côté, *arica-* de l'autre. Mais le situationnel (dit « présent ») est formé, en grec, à partir de **leik^{w}o-*, d'où *leípō* « je laisse », en sanskrit à partir de **li-na-*

k^w, d'où *rinakti* « il laisse ». En latin, où il ne s'oppose qu'au parfait (*līquī*), *linquo* dérive de *li-n-k^{w}o- qui rappelle un peu la forme sanskrite, mais avec un *-n-* au lieu de *-na-* et un suffixe *-o-* final. A l'origine des formes gotiques, on a, comme en grec, *leik^{w}o. Mais, comme en latin, cette forme s'oppose seulement à un parfait, qui, comme le grec *lé-loip-a*, provient de *loik^w-.

Nous avons vu que le sens de certains verbes, *être*, par exemple, pouvait rendre difficile la découverte de partenaires adéquats. On y est souvent arrivé, comme dans le cas de *être* en latin, en combinant des verbes d'origines très diverses. C'est ce qu'on désigne comme la supplétion. En grec, par exemple, la notion de « manger » s'exprime, au parfait (*édēda*), au présent (*esthiō*) et au futur (*édomai*) par des dérivés de la racine *ed-* représentée en latin par *edo*, en anglais par *eat* ; à l'aoriste, on a *phágo-mai*, sans doute, au départ, « avaler un morceau ». En latin, où l'expression de l'événementiel se faisait au moyen du parfait, il a fallu procurer cette forme à *fero* « porter » qui n'en comportait pas. On a donc emprunté celui de *tollo* « soulever », *tulī*, qui vaut donc pour l'un et l'autre verbes.

On peut donc, en ce qui concerne le processus de « conjugaison », répartir les langues indo-européennes en trois groupes. Un premier qui ne l'a pas connu et qui est représenté par le hittite. Un second, avec les langues de l'ouest, germaniques, celtiques, italiques, où la conjugaison se limite à deux formes, 1. un parfait qui évolue vers un passé en s'annexant éventuellement certaines formes de l'ancien « nom d'action » et qui s'oppose à 2. un situationnel dit « présent » même s'il présente des temps divers. Le troisième groupe avec, essentiellement, le grec et l'indo-iranien, « conjugue » les trois types distincts : parfait, événementiel et situationnel.

Les temps.

Jusqu'ici, nous n'avons pas postulé l'existence de temps verbaux à date ancienne. Si nous n'étions pas influencés, dans notre jugement, par nos langues contemporaines, nous verrions que rien n'oblige à replacer toute action dans une tranche de temps coïncidant avec l'instant où l'on parle, ou antérieure ou postérieure à cet instant. Pourquoi, lorsque cette prise de position est nécessaire à l'expression de son expérience, ne pas utiliser quelque adverbe adéquat ? Notons d'ailleurs que dans les langues indo-européennes contemporaines et un peu partout ailleurs, on dispose d'une forme verbale qu'on désigne comme le présent, mais qui implique simplement le procès dans son déroulement et qu'on peut parfaitement employer en référence à une action passée (présent « historique » ou de « narration ») ou à venir : *demain, je pars pour New York*, angl. *when he comes (I shall...)*

« lorsqu'il viendra... je... -rai ». Il est clair que le temps du discours ne peut pas être considéré comme une modalité du procès. Ce qui l'est, c'est, par exemple, ce que nous avons rencontré ci-dessus quand nous avons traité des divers types de situation dynamique : l'action qui débute, celle qui se poursuit, celle qui aboutit, auxquelles on peut ajouter celle qui se répète, sans préjudice de l'action conçue dans son aboutissement (événementiel) et de l'action accomplie, tout ce qu'on désigne comme les aspects du verbe.

Comment l'expression du temps finit par se greffer sur celle des aspects, ceci a déjà été suggéré dans ce qui précède. Quand on entend *j'ai mangé* de la bouche de quelqu'un qui veut simplement faire comprendre qu'il n'a plus faim, on saisit parfaitement le message, mais rien n'empêche qu'on imagine la personne en cause prenant effectivement de la nourriture antérieurement à la production de son énoncé. Un accompli a toute chance d'être vite utilisé, sporadiquement peut-être, mais effectivement, comme un passé.

Dans le cas de l'événementiel, il n'est même pas nécessaire de supposer un processus évolutif qui mènera d'un aspect accompli à un temps passé. Logiquement, et on l'a relevé depuis longtemps, on ne peut pas signaler un événement au présent, parce que le temps nécessaire à en prendre connaissance et à l'expliciter fait qu'il va nécessairement appartenir au passé : on ne peut pas dire « la boule touche le plafond » parce qu'au moment où on le dit, la boule est sur le plancher. Ceci se vérifie en grec où l'aoriste exprime toujours un passé, sauf dans les contextes grammaticaux où le temps ne joue plus : à l'infinitif, l'aoriste *phágein* veut dire « avaler un morceau » sans référence au temps.

Les aspects verbaux ont donc toujours des implications temporelles qui s'imposent plus ou moins aux dépens des modalités d'action qu'ils sont censés exprimer. Mais le premier temps qui s'établit est toujours le passé qui seul, avec le présent, peut ne pas être, dans un sens, hypothétique. Le futur proprement dit est plus rare, parce que, étant toujours, dans l'esprit, mêlé à des désirs ou à des obligations, ce sont ces désirs et ces obligations qui s'imposent en priorité à l'attention. C'est ce qu'on constate, en anglais, avec *will* et *shall*, au départ « vouloir » et « devoir », dont le premier surtout est franchement futurisé en comparaison avec *want to* et *have to* où désir et obligation restent seuls pertinents. En latin, le futur dérive soit de désidératifs, soit, comme en allemand avec l'auxiliaire *werden*, de l'adjonction d'un « devenir », avec les formes en *-b-* (**bhew*) : *amābo* « j'aimerai », *uīdēbo* « je verrai ».

Les modes.

On entrevoit par l'examen de ces dernières formes combien peut être instable la frontière entre les temps et les modes. Si on laisse de côté le mode zéro qu'on nomme l'indicatif, il n'y a, parmi les modes dits personnels, c'est-à-dire proprement verbaux, que l'impératif qui se retrouve partout, en indo-européen, parfois avec un paradigme complet de neuf personnes, duel compris, comme en sanskrit, ou de six comme en hittite, ailleurs limité, comme en français, à l'adresse de personnes en situation, locuteur excepté, c'est-à-dire « toi », « nous » et « vous ». La forme la plus permanente est la 2e personne du singulier, « toi », souvent représenté par le radical nu : *donne !, va !* Elle peut être complétée par des auxiliaires, comme en anglais, dans *let him go !, let them do !,* ou renforcée par des particules adverbiales diverses, comme *donc* dans *dis-donc !* Il y a, par exemple, un élément *-u* qu'on retrouve aux 3es personnes du singulier et du pluriel, aussi bien en hittite qu'en sanskrit. Mais de tels éléments sont souvent instables et susceptibles d'apparaître çà ou là. Un cas intéressant est celui de ce qu'on désigne en latin comme l'impératif futur en *-tō : crās peti-tō* « demande demain ! » On a suggéré que ce *-tō* n'est pas autre chose que l'ablatif du démonstratif, donc *tōd,* avec la valeur de « à partir de ce (moment-)ci », « désormais ».

Des autres modes attestés dans les langues indo-européennes, c'est l'optatif qui a eu la plus vaste diffusion, avec une marque très caractéristique *-y-*. Ce qu'on désigne, ici et là, comme un subjonctif paraît être d'origine assez hétéroclite : le subjonctif français du verbe *être, soit,* remonte à un optatif, mais celui de *savoir, sache,* à un désidératif. Le terme même de subjonctif suggère qu'il ne s'emploierait que dans des propositions subordonnées, ce qui est à peu près le cas en français et qui restreint beaucoup son pouvoir informatif. Ce dernier est, en français, très faible, car le subjonctif y apparaît surtout dans des contextes où le choix d'un autre mode n'est pas possible, et là où il pourrait, à lui seul, marquer une différence *(je cherche quelqu'un qui sache... je cherche quelqu'un qui sait...),* on ne peut l'employer avec sécurité qu'avec des verbes irréguliers ou des personnes (1re et 2e du pluriel) où l'on distingue encore entre indicatif et subjonctif.

Le présent hic et nunc.

On a vu ci-dessus que l'expression du passé au moyen de formes verbales résulte un peu partout, et cela a été le cas en indo-européen, directement ou indirectement, de l'expression des aspects événementiel ou accompli. Mais il est clair qu'assez tôt l'opposition du passé au

présent s'est imposée et a été marquée de deux façons différentes. La première a résulté de l'extension d'un élément *i*, indiquant positivement l'existence de quelque chose ou de quelqu'un dans l'espace ou l'instant où l'on parle. On a cité ci-dessus une forme grecque comme *houtos-í*, renforcement de *hoûtos* « celui-ci ». Cet *i* a été utilisé pour marquer ce qu'on peut appeler un présent *hic et nunc*, un présent positivement marqué comme tel. Ce qui montre combien il s'est longtemps agi d'un élément indépendant, c'est que sa répartition varie beaucoup de langue à langue. Elle est maximale en hittite où le *-i* caractérise les personnes de tous les présents de n'importe quelle origine partout où ils s'opposent à un passé, donc à l'indicatif, mais non à l'impératif. Ailleurs, il n'apparaît ou ne se manifeste indirectement qu'au singulier et à la 3[e] personne du pluriel ; en sanskrit, par exemple, dans 1 s. *bhavāmi*, 2.s. *bhavasi*, 3.s. *bhavati*, 3.pl. *bhavanti*. En grec, on le trouve à 1 s. dans les verbes à radical consonantique : *títhēmi* « je place » (< *-$d^h eH_1$-), mais il est absent après les radicaux vocaliques : *luō* « je dételle ». A la voix moyenne, sur laquelle nous reviendrons longuement ci-après, on trouve le *-i* intégré à la voyelle finale de la désinence en une diphtongue *-ai*. En grec, avec limitation aux mêmes personnes qu'à la voix active : *tithemai* « je pose pour moi, sur moi », *títhesai*, *títhetai*, *títhentai*, mais, en sanskrit, également à 1 pl. *bhavāmahe* et 2 pl. *bhavadhvē* (avec *-ē* < **ai*). Au parfait qui est, bien entendu, au départ, un présent, le *i* s'est ajouté au *-a* final (< *-H_2e) pour donner **-ai* qui apparaît en latin comme *-ī* dans *uixī* « j'ai vécu ». En grec et en sanskrit, *-ai* (d'où en sanskrit *-ē*) est, dans ce cas, limité à la voix moyenne.

L'augment.

Cette présence de *-i* opposée à son absence s'exprime traditionnellement sous la forme d'une opposition entre désinences dites « primaires » avec *i* et désinences dites « secondaires » sans *-i*. Le choix de ces épithètes marque bien que le présent en *-i*, qui était, au départ, marqué comme tel, était devenu un « présent » ordinaire, c'est-à-dire la forme verbale non marquée quant au temps. Cette évolution a, naturellement, amené à sentir le passé comme une forme sémantiquement marquée, ce qui incitait à lui trouver une marque formelle. C'est ce qui a fait apparaître la seconde façon d'indiquer l'opposition du présent au passé sous la forme de ce qu'on appelle l'augment. Il s'agit d'une particule adverbiale *H_1e qui précède, en sanskrit et en grec, les formes de sens passé : skrt. *a-bhavam* « je devenais », *a-bhuvam* « je devins », grec *é-lipon* « je laissai », *e-títhēn* « je plaçais » et, avec le même sens, une ancienne forme de parfait passée à l'aoriste, *é-thē-k-a*. Le parfait, un présent, n'a pas l'augment : *(w)oída* « j'ai vu, je sais »,

lé-luka « j'ai dételé (c'est fait) », mais le « plus-que-parfait », passé du parfait, a l'augment : *e-le-lúkē* « j'avais dételé (c'était fait) ». Mais, plus encore que dans le cas du *i* de présent, les locuteurs sont restés longtemps libres de l'employer ou non. Il est intéressant de noter que l'augment n'est attesté, outre les deux langues déjà citées, qu'en arménien et en phrygien, ce qui suggère qu'il est apparu et s'est diffusé dans un groupement de tribus indo-européennes qu'on pourrait situer, entre le quatrième et le troisième millénaire, à l'ouest de la Mer Noire. Ceci laisserait, vers le nord-est, les ancêtres des Baltes et des Slaves, vers le nord-ouest ceux des Occidentaux, celtes, germains et italiques, et naturellement, en Asie mineure, les Anatoliens.

Les modes non personnels.

Lorsque les anciens « noms » d'action et d'auteur se sont singularisés comme verbes, avec des flexions personnelles, le besoin s'est fait sentir de formes proprement nominales désignant les actions en elles-mêmes ou leurs participants actifs ou passifs, d'où apparition des formes désignées comme l'infinitif et les participes. Les infinitifs varient beaucoup d'une branche à une autre : en grec, on a des formes diverses, la plus productive étant en *-n* : *légein* « parler » ; en latin s'est imposée une forme en *-s-* passée régulièrement à *-r-* : *legere* « lire ». Les participes sont plus uniformes. Les radicaux en *-nt-*, dits « participes présents actifs », sont attestés du sanskrit (*bharant-* « apportant ») au latin (*ferent-* même sens) en passant par le grec (*pheront-*), le slave (*nesǫšt-* < **nesonti-* même sens) et le germanique (*berend-*). Les « participes présents » moyens ou passifs en *-meno-*, *-mono-* du grec et du sanskrit, sont attestés ailleurs comme adjectifs, par exemple, dans le latin *alu-mnu-s* « élève », très précisément « celui qu'on élève », de *alo* « je nourris, j'élève », ou comme substantif : *fēmina* « celle qui donne à téter », d'une racine *d^hē(i)* qu'on retrouve dans *fēlāre* « téter » et dans *fīlius* « fils » (cf., ci-dessous, p. 233). De même le participe parfait en *-wes-*, *-wos-*, alternant avec *-wet-*, *-wot-*, du grec et du sanskrit, est attesté dans le got. *wit-wōþs* « témoin » qui assume l'ancienne valeur du simple **woide*, grec *(w)oíde* « il sait (parce qu'il a vu) ».

L'élément *-to-*, qui est à la source des participes parfaits du latin et des langues romanes, n'avait pas au départ un sens nécessairement passif : *pō-tu-s* « qui a bu », *taci-tu-s* « qui n'a rien dit ». C'était un adjectif formé sur le « nom d'action » en *-t* : **peH_3-t* (> **pō-t*) « action de boire », **peH_3-t-o* (> **pō-t-o*) « concerné par l'action de boire ».

Les voix.

Lorsqu'on parle de voix, ou, plus savamment, de diathèses, on pense, en priorité, dans les langues modernes, à l'actif et au passif : *il tue* ~ *il est tué*. En français, il n'y a pas de véritable passif hors de ce qu'on appelle le « participe passé ». On en dérive des constructions ambiguës, comme *il est tué*, qui a deux sens très différents : 1. « Il est là, qui a été mis à mort », ce qui est un parfait, et « on est en train de le mettre à mort », ce qui est un situationnel dynamique. L'allemand a un passif, périphrastique sans doute, mais bien caractérisé, avec *er wird getötet* dont la meilleure traduction est « on le tue ». Les langues scandinaves ont une forme verbale distincte, dan. *han dræbes*, même sens, à côté du calque de l'allemand *han bliver dræbt*. L'équivalent italien *viene ucciso* n'est pas ambigu. L'anglais peut éviter la confusion au moyen de l'auxiliaire *get : he got killed* est incontestablement un passif. Les constructions réfléchies sont souvent équivalentes au passif : l'équivalent de l'angl. *English (is) spoken*, qui est en français *on parle anglais*, est en italien *si parla inglese* et, en espagnol, *se habla inglés*. Tout ceci témoigne de la difficulté de se procurer un passif bien marqué comme tel. Ce dont on a besoin réellement, à côté de l'actif, c'est d'une forme où l'auteur de l'action ne soit pas explicité, soit parce qu'on ne le connaît pas, soit parce qu'on n'a pas besoin de le mentionner ou pas intérêt à le faire. Un passif complet avec expression de l'agent, comme dans *le portail a été fermé par le jardinier*, est un luxe d'une utilité discutable là où l'on peut dire *le jardinier a fermé le portail*.

Tout cela explique que l'on ne puisse pas reconstruire un passif pour l'indo-européen commun. Ce que l'on trouve dans les langues les plus anciennes, du hittite au grec et au sanskrit, c'est ce qu'on appelle, curieusement, le « moyen ». La valeur du moyen couvre à peu près celle du réfléchi et, de ce fait, il peut servir à exprimer l'équivalent du passif. Mais il s'emploie également lorsque la participation de l'agent à l'action n'est pas ponctuelle et extérieure. Si je fends du bois, une fois le coup de hache asséné, le changement effectué ne me touche plus. Si en revanche on suit quelqu'un, on participe à l'action aussi longtemps que l'objet de la filature. Si j'imite quelqu'un, je suis beaucoup plus directement impliqué que celui que j'imite. Ceci explique que « suivre » a normalement en grec et automatiquement en latin la forme d'un « moyen » : grec *hépomai*, lat. *sequor*. Il en va de même pour « imiter » : grec *miméomai*, lat. *imitor*, et pour « se nourrir » : grec *tréphomai*, lat. *uescor*, etc.

Comme on vient de le constater avec *sequor* et *imitor*, il y a des verbes qui n'apparaissent qu'au « moyen ». Ils ont reçu traditionnellement le nom de « déponents ».

Le « moyen » est marqué comme tel et distingué de l'autre voix, ou absence de voix (désignée comme l' « actif » comme si elle s'opposait à un passif), de façons assez diverses selon les langues : on vient de voir qu'à partir d'un radical **sek*^{w}*o-* « suivre », le grec marque la 1re personne du singulier du « moyen » par *-mai*, alors que le latin utilise le *-r*. La finale *-mai* comporte évidemment le *-m* de 1.s. et le *-i* de présent. L'équivalent de *-mai* au passé est *-mo*, sans *-i*, bien entendu. Le *-a-* de *-mai*, au lieu de *-o* pose un problème. On pense à une influence du *-a* (< *-H_2e) du parfait, et cette influence s'explique assez bien lorsqu'on se rappelle que la construction à partir d'un nom d'auteur, qui a donné le parfait, ne suppose pas une action et l'objet de cette action. Le « moyen », de son côté, confond le sujet et l'objet dans lat. *uescor* « je me nourris », le sujet est moi-même et l'objet également ; ce de quoi je me nourris est une circonstance extérieure marquée en latin par un ablatif correspondant au français « de » ; le grec *hépomai* est normalement suivi du datif comme son équivalent allemand *folgen*. C'est cette vocation intransitive initiale, commune au parfait et au moyen, qui justifie donc *-mai* au lieu du *-moi* que laisserait attendre le *-mo* du passé. Les désinences « moyennes » pourraient donc être, à l'origine, celles de l'actif plus *-o*.

Le *-r* du latin provient de l'impersonnel que nous avons rencontré ci-dessus lorsque nous traitions du problème de l'alternance *-r*/*n-*. Soit la racine *ei-*/*i-* « aller » encore attestée dans le français *ira*. Le « nom d'action » correspondant est **ei-t*. Il s'y ajoute un suffixe **-en* qui devient *-er* à la finale lorsque **nt* final passe à *-n*. Accentué sur le suffixe, **ei-t-ér* donne **i-t-ér*, d'où le latin *iter* « chemin que l'on fait ». Accentué sur la racine, on a **éi-t-r̥*, d'où *itur* « on va », en quelque sorte « (il y a) cheminement ». Comme le « nom d'action » en *-t-* est interprété comme la troisième personne du singulier, les formes en *-t-ur-*, lorsqu'elles se rattachent à un transitif comme **leg-o-* « lire », peuvent être accompagnées d'un objet, ce qu'on lit, par exemple, et cet unique participant de l'impersonnel prend la forme du sujet : *librum legitur* « on lit le livre » qui est attesté, passe à *liber legitur* « le livre est lu ». La forme en *-tur* devenue un passif, sa marque caractéristique *-(u)r* va s'étendre à d'autres personnes, comme la 3e du pluriel : *legunt-ur* « ils sont lus », la première du singulier : *amo-r* « je suis aimé » et du pluriel : *amāmu-r* « nous sommes aimés ». Ce nouveau passif n'est pas dissocié du « moyen », d'où les formes des « déponents » *sequitur, sequuntur, sequor, sequimur* qui gardent leur valeur « moyenne » et un sujet qui est l'agent. A la deuxième personne du singulier, pour les passifs aussi bien que les « déponents », on garde l'ancienne forme « moyenne » *amāre* (< **amā-s-e*) ; lorsque le *-s-* intervocalique est devenu régulièrement *-r-*, on a rajouté un *-s*, caractéristique de cette deuxième personne, d'où **amāre-s* > *amāri-s*.

Les formes en *-r* sont très bien représentées en celtique avec, comme en latin, des passifs et des déponents, mais résultant d'évolutions en partie différentes : en vieil-irlandais, par exemple, seuls les déponents présentent des désinences personnelles distinctes. Le passif a gardé une forme d'impersonnel en **-tr̥* (< *-thar*) à toutes les personnes, celles-ci étant indiquées au moyen de pronoms objets infixés. A partir de *caraid* « il aime », on forme un passif *-carthar* « on (l')aime », avec *no-m-charthar* « on m'aime », *no-t-charthar* « on t'aime », où *no-* ne sert qu'à introduire les pronoms objets *-m-*, *-t-*, etc.

En tokharien, les formes en *-r* sont intégrées à la flexion du « moyen », avec des formes personnelles variables comme 1 s. *-mār(ə)*, 2. s. *-tār(ə)*, 1. pl. *-mtrə*, 3 pl. *-ntrə*. Il s'agit bien, comme en latin, d'une extension à partir d'un impersonnel en *-tr̥* représenté par 3 s. *-tər(ə)*. En hittite, *-r-*, suivi de *-i* ou de *-u* selon le mode, s'ajoute, comme en latin, aux désinences personnelles, comme dans *ya-ha-r-i*, de *ya-* « aller » et *-h(a)-*, marque de première personne. C'est une des marques du « moyen ».

La forme sans *-t-* qu'on trouve en hittite, en celtique (irl. *berir* « on porte »), ainsi que dans l'ombrien *ferar* « qu'il soit porté » en face de l'équivalent latin *ferātur*, pose quelques problèmes. On pourrait y voir une combinaison de *-r* (< **-n*) non plus avec le « nom d'action » en *-t-*, mais avec le radical du parfait, l'ancien « nom d'auteur » qui est en voyelle (**woide*). Si, pour comprendre les formations, on peut traduire *-r* par « il y a », une telle combinaison voudrait dire quelque chose comme « il y a quelqu'un qui a fait », d'où « on a fait ». Mais il se trouve que les « moyens-passifs » en *-r* ne sont pas représentés au parfait : en latin, le parfait de *sequor* est une forme composée, *secutus sum ;* en irlandais, celui de *caraid*, qui est *-carthar* au présent, est *-carad* sans *-r*. En tokharien et en hittite, le parfait n'est pas représenté comme tel. Toutefois une finale en *-r* est bien attestée au parfait, en sanskrit, dans les troisièmes personnes du pluriel, comme l'actif *ca-kr-ur* « ils ont fait » et le moyen *ca-kr-ire* « ils se sont fait ». De là, elle s'est étendue aux temps du passé, comme l'imparfait *a-da-dh-ur* « ils plaçaient » ou l'aoriste *a-d-ur* « ils donnaient ». En latin, une forme de 3 pl. parfait comme *dixere*, à côté de *dixerunt* atteste *-r-* avec la même valeur. En hittite, les troisièmes personnes du pluriel du prétérit sont régulièrement en *-er : w-er* « ils vinrent », *pahs-er* « ils protégeaient ». On peut donc penser que des impersonnels en *-r* formés sur le radical du parfait ont pu prendre valeur de 3 pl. Il est fréquent qu'on emploie les 3 pl. avec valeur impersonnelle, comme en anglais *they say* qui peut valoir « on dit ».

LE NEUTRE

On pourrait s'étonner que le neutre, classé comme un genre, n'ait été mentionné ci-dessus que de façon épisodique lorsqu'on s'est occupé des formes nominales. C'est que si, finalement, il participe à l'accord, comme le féminin dont c'est la seule caractéristique fondamentale, le neutre ne s'oppose au départ au non-neutre, que du fait que les nominaux de ce genre ne distinguent pas la forme du sujet de celle de l'objet : le latin *animal*, qui est un neutre, désigne aussi bien l'animal qu'on voit (objet) que celui qui court (sujet) ; de même le pluriel *animalia* vaut comme sujet et comme objet sans distinction de forme. Pour comprendre comment de tels nominaux peuvent coexister avec d'autres qui distinguent formellement sujet et objet, comme les masculins lat. *dominu-s* ~ *dominu-m* « maître », *cīui-s* ~ *cīue-m* « citoyen » ou les féminins *nuru-s* ~ *nuru-m* « belle-fille », *fēmina* ~ *fēmina-m* « femelle », il est indispensable de revenir sur la façon dont s'est constituée la « conjugaison » par le rapprochement de formes verbales, ou « préverbales », dont **les rapports avec les éléments du contexte** pouvaient varier du tout au tout. Rappelons que ces formes, dans un état de langue où verbes et noms ne s'opposent pas, ne sont proprement ni verbes, ni noms. Mais il faut envisager, pour ces unités indifférenciées, des relations au contexte que nous saisissons mieux si nous utilisons des traductions nominales. C'est pour bien marquer le caractère pragmatique de l'emploi de « nom » dans ce cas que nous conservons les guillemets dans « nom d'action », « nom d'auteur », etc.

Reprenons notre racine $*g^{hw}$*en*, et son nom d'action $*g^{hw}$*en-t-* (skrt. *hant-i* « (il) frappe »), d'où, avec une particule de première personne, $*g^{hw}$*en-t-mo* > $*g^{hw}$*en-mo*. Une telle forme n'est pas attestée comme telle, parce que la racine y est, à tout moment, susceptible de voir son intégrité phonique affectée par la nature de l'élément possessif suivant, $*g^{hw}$*en-mo* passant à $*g^{hw}$*em-mo* avec *-m-* en face de g^{hw}*en-t* qui garde l'*-n-*. C'est pourquoi ces formes seront finalement remplacées par des dérivés plus stables, comme $*g^{hw}$*en-yo-* qui a abouti au grec *theínō* « je frappe ».

La forme $*g^{hw}$*en-mo* vaut, à peu près, « frappe de moi », « mon action de frapper (à mort) ». La relation entre l'action et le participant représenté par *-mo* est celle qu'on marquera au moyen de la particule $*$*es* si le participant est représenté par un substantif. Avec $*$*reg* « chef, roi » on aura $*g^{hw}$*en-t reg es* « (il y a) frappe qui vient du chef », « ... frappe par le chef ». Si la mention de la frappe est accompagnée de la mention de la personne ou de l'objet qui la subit, le rapport syntaxique d'une forme à l'autre sera indiqué simplement par l'antéposition de la

seconde à la première. C'est ce rapport et sa marque qu'on retrouve dans des composés du type de lat. *agricola* « agriculteur » où *-col-* implique l'action de cultiver et *agri-* (< *agro-*) représente le radical nu et a la valeur d'un génitif (génitif$_1$ de notre tableau, p. 195) Soit, par exemple, *g^hwer*, qui a dû désigner l'animal sauvage (cf. grec *thēr*, v. slave *zvěrĭ*, de même sens, et lat. *feru-s* « sauvage », de *g^hwer-o-*). On peut reconstruire un énoncé comme **reg es g^hwer g^{hw}en-t-i*, avec un *-i* de présent *hic et nunc* et la valeur « (il y a) par le chef — (de) l'animal — mise à mort — en cet instant ». Le syntagme **reg es* évoluera en **reges* > *rēks* qui est, avec x dans la graphie pour *-ks*, la forme latine pour « roi ».

Les noms d'état.

A côté des « noms d'action » en *-t*, il faut poser des « noms d'état ». Il doit y en avoir où le radical s'identifie à une racine, ainsi *H_1es* « existence ». Mais on en forme aussi par dérivation, par exemple, au moyen d'un suffixe **-eH_1* qui donnera *-ē*. Soit **lat-eH_1* « latence », « présence cachée » ou **lub^h-eH_1* « préférence, appréciation » qui apparaîtront plus tard, en latin, comme *late-t*, *lube-t* ou *libe-t*, avec, lorsque la « conjugaison » aura été réalisée, assimilation formelle aux « noms d'actions », d'où le *-t* final de 3e personne. Leur rapport syntaxique avec le contexte est celui que nous avons rencontré ci-dessus dans *agricola* « agriculteur » où le premier élément a la valeur d'un génitif, d'où **reg H_1es* « (il y a) — d'(un) chef — existence ». De même **reg lat-eH_1* « (il y a) — du chef — présence cachée ». Si doit intervenir un autre participant, ce sera sous forme indirecte, comme un complément d'attribution marqué par la particule *ey* : **reg ey... lubeH_1* « (il y a) attribuée au chef... préférence ». On a encore en latin *mihi libet* avec, au datif, la désignation de la personne qui éprouve la préférence ou qui apprécie. Noter le *me* du français *ça me plaît*.

De la racine **g^{hw}en* « frappe », on dérive un « adjectif » au moyen du suffixe *e/o*, d'où **g^{hw}ene* « (celui) de la frappe ». On distinguera ultérieurement par l'accent entre **g^{hw}ené* > **g^{hw}oné* « meurtrier », « auteur de la frappe », et **g^{hw}éne* > **g^{hw}éno* et, par analogie, **g^{hw}óno-* « le résultat de la frappe », « le meurtre », qui donne régulièrement, en grec, *phóno-s*. Comme nom d'auteur, **g^{hw}oné*, « celui qui a frappé », va pouvoir être accolé à la désignation d'un individu, **reg* « chef », par exemple, ou un personnel comme **H_2e* « moi » dans un rapport d'identification. Cette relation, qui est presque celle des deux termes d'une équation, doit être conçue comme celle qui sera marquée plus tard par la copule. Elle n'indique donc, en aucune façon, subordination d'un élément à l'autre et elle n'est pas formellement marquée. On aura donc **reg g^{hw}oné* « chef = meurtrier », « (c'est) le

chef (qui a) frappé », sans addition. Si la personne en cause est « la première », « moi », le pronom n'aura pas la forme quasi possessive *-mo,* mais celle de l'asyntaxique H_2e. La combinaison du « nom d'auteur » avec un personnel asyntaxique donnera ce qu'on désigne comme le parfait.

Dans l'état de langue dont nous venons d'esquisser les traits syntaxiques fondamentaux, la forme nue des radicaux substantivaux (du type **reg*) et les formes asyntaxiques des pronoms personnels (du type $-H_2e$) se combinent sans restriction avec n'importe lequel des « noms » qui sont à la source de nos verbes pour désigner 1) le patient, avec les « noms d'action » (« la bête » dans « le chef tue la bête »), 2) le participant central pour les « noms d'état » (« le chef » dans « présence cachée du chef »), 3) , avec ce qui deviendra le parfait, aussi bien l'auteur de l'action qu'un patient éventuel affecté par l'action, c'est-à-dire dans « le chef = celui qui a tué la bête », aussi bien « chef » que « bête ». Cette forme est ce qu'on désigne comme un **absolutif.** Dans les contextes non vocatifs et non proprement nominatifs, où elle assume une valeur réellement syntaxique, on peut dire qu'elle est le premier déterminant du noyau prédicatif, même avec un « nom d'action » comme $*g^{hw}en\text{-}t$ qui, au départ, n'implique obligatoirement aucun agent et vaut simplement pour « (il y a) frappe ». Le déterminant attendu dans ce cas est le patient, par exemple $*g^{h}wer$ $g^{hw}en\text{-}t$ « (il y a) frappe de l'animal sauvage ».

Du nom d'action au verbe.

La situation va changer lorsque la forme en *-t* va acquérir un statut de 3ᵉ personne, lorsque, par exemple, sur le modèle de 1. $*g^{hw}en\text{-}mo$, 2. $*g^{hw}en\text{-}so$, formes qui donneront plus tard le « moyen », on forme une 3. $*g^{hw}en\text{-}to$ et des présents *hic et nunc,* « actif », $*g^{hw}en\text{-}t\text{-}i$ et « moyen », $*g^{hw}en\text{-}to\text{-}i$. Dès lors, l'expression de l'agent va devenir automatique, obligatoire, pour tous les « noms d'action » qui prennent, avec leur flexion personnelle, ce que nous pouvons appeler un statut verbal. Pour ceux qui deviennent désormais verbes d'action, s'impose un complément obligatoire qui va, du fait de sa présence permanente, tendre à assumer la valeur de thème du discours, ou, si l'on veut, de sujet, et à réduire le patient au rôle d'annexe facultative : le noyau de l'énoncé n'est plus « frappe de l'objet », mais bien « quelqu'un frappe ».

La flexion personnelle au moyen des anciens possessifs va s'imposer aux « noms d'état », devenant, eux aussi, des verbes, donc, non plus $*H_1es$ « (il y a) existence », mais $*H_1es\text{-}t\text{-}i$, avec valeur de « il existe », $*H_1es\text{-}m\text{-}i$ « j'existe », etc. Cependant, là-même où le *-t-* s'impose, certains verbes d'état n'étendent pas la flexion à d'autres personnes

que la troisième et restent donc ce qu'on appelle les impersonnels, comme en latin *libet* « il plaît » qui cependant peut prendre comme sujet *un pronom neutre : quod libet* « ce qui plaît », la personne en cause continuant à figurer au datif : *quod mihi libet* « ce qui me plaît ».

Les substantifs qui ne désignent pas des entités actives n'ont guère d'occasion d'apparaître comme sujets des verbes d'action. Ils n'auront donc aucune occasion de se voir affixer l'élément *-s* (< **es*) qui marque l'agent. Avec les verbes d'état, l'unique participant peut être aussi bien une entité active, par exemple, « l'homme » dans « l'homme paraît fatigué », qu'une entité passive, comme « le trésor » dans « le trésor est caché ». Cet unique participant se présentait, au départ, sous la forme du radical nu, mais l'extension aux verbes d'état de la flexion personnelle des verbes d'action va entraîner l'utilisation de formes substantivales en *-s* parallèles aux désinences personnelles. Ceci ne se produira toutefois que *dans la mesure* où ces formes sont bien ancrées dans la mémoire du fait de leur fréquence dans les constructions avec les verbes d'action. On conservera donc, pour les entités généralement passives, les radicaux nus, les enfants qui apprennent la langue n'ayant jamais l'occasion de les entendre avec un *-s* affixé. Soit, par exemple, la mer, lat. *mare*, qui, on se le rappelle, désigne, au départ, une étendue d'eau calme. Elle s'emploiera fréquemment avec valeur ablative (« il vient de la mer ») indiquée par la particule **és*, devenue ultérieurement la désinence *-és*. Cette combinaison sera ultérieurement utilisée comme génitif au lieu de la simple forme nue juxtaposée. Mais, comme cette eau calme est soit l'endroit où quelque chose se passe (locatif), soit ce au sujet de quoi on peut dire quelque chose, mais n'est pratiquement jamais présentée comme l'initiatrice d'aucune action, elle n'apparaîtra guère munie de la marque *-s* de l'agent. Les locuteurs ne seront donc jamais tentés de remplacer le radical **mori*, ou **mari*, qu'on emploie, de fondation, avec les « noms d'état », puis avec les verbes qu'ils sont devenus, par un **mori-s* qu'on n'entend jamais ailleurs. Soit encore **kreid^{h}ro*, l'instrument qui trie, le crible, lat. *cribru-m*, angl. *riddle* (< *hridder*). Avec un verbe d'action, on le voit très bien utilisé à l'instrumental ou comme objet. Rien n'a pu empêcher un locuteur doué d'imagination de le présenter comme l'initiateur d'une action, mais **kreid^{h}ro-s* avec *-s*, restait une plaisanterie isolée, peu susceptible d'influencer le comportement des générations à venir. La forme nue **kreid^{h}ro* était celle qu'on utilisait pour le participant unique des verbes d'état et en rapport avec un attribut substantival ou adjectival : « le crible, c'est un instrument qui... », « ce crible est inutilisable ». Dire que lat. *cribru-m* est « du neutre », c'est exactement dire que sa forme de base, tardivement pourvue d'un *-m* comme nous le verrons plus loin, s'emploie aussi bien pour l'objet

d'un verbe d'action que pour le participant unique d'un verbe d'état, c'est-à-dire pour ce qu'on appelle un sujet.

Dans le cas du parfait, la flexion traditionnelle, au moyen d'anciennes formes pronominales asyntaxiques du type 1. *-H_2e*, n'a pas été affectée, au moins aux 1re et 2e personnes du singulier. Elle était là de fondation, utilisée constamment et certainement très tôt par les enfants. Ceci a pu retarder l'extension au parfait de l'emploi de la forme d'agent en *-s* comme équivalente d'un élément pronominal qui, ici, avait la forme zéro. En d'autres termes, en face d'une construction active où le *-t-* de 3e personne du verbe allait de pair avec le *-s* du sujet, a dû se maintenir quelque temps une construction de parfait où au zéro de la 3e personne du verbe correspondait le radical nu, à désinence zéro, pour ce qui était en voie de devenir un sujet. Le déterminant éventuel du « nom d'auteur » était un radical nu de valeur génitivale. Dans un premier temps, il ne va pas changer de forme, mais graduellement s'assimiler, quant à sa valeur, à l'objet des verbes d'action.

Lorsque s'établit la « conjugaison », on a donc en présence — outre la proposition nominale correspondant à ce qui sera, plus tard, deux nominaux réunis par la copule — trois types fondamentaux de construction :

1) la construction des verbes d'action avec objet (construction transitive) ;
2) la construction des verbes d'état qui va se confondre avec celle des verbes d'action sans objet ;
3) la construction des parfaits (transitive ou intransitive).

On peut les visualiser comme suit, avec un verbe V à la 3e personne et une première colonne où figure ce qui devient le sujet, une seconde avec ce qui devient l'objet, une troisième avec le verbe :

1.	N-s	N	V-t
2.	{ N-s { N		V-t
3.	N	N	V

C'est la forme de l'objet qui doit maintenant retenir notre attention. Là où il figure, chez les transitifs des types 1 et 3, il garde la forme non marquée du radical nu. Au type 1, il se distingue parfaitement du sujet caractérisé par *-s*. Au type 3, en revanche, la confusion est possible. La position de l'objet immédiatement avant le verbe, telle qu'elle figure dans le schéma qui précède, n'assume certainement pas une véritable pertinence, surtout, sans doute, dans le type 1, mais également, par extension, dans 3. Il y aura donc des cas où il sera nécessaire d'établir

une distinction formelle entre sujet et objet là où elle n'existe pas, et c'est là qu'intervient l'utilisation du nouvel allatif en *-m* pour marquer l'objet. C'est, on l'a vu, ce qui s'est produit en espagnol où l'allatif, marqué par la préposition *a*, s'emploie là où le sens du terme en fonction objet pourrait suggérer un rôle actif pour son référent. C'est ce qu'on constate en indo-européen. Seuls les non-neutres, susceptibles d'être conçus et employés comme agents, y reçoivent le *-m* d'allatif. Des neutres comme lat. *mare, pecu* « bétail », *animal, iecur* « foie » apparaissent, comme objet, sous la forme du radical nu.

Les neutres en -o-m.

L'exception, et elle est d'importance, est représentée par les formations adjectivales en *-e/o* qui, elles, reçoivent le *-m* de l'accusatif. Ce traitement particulier s'explique bien. Soit, à partir de **nu*, **new* « maintenant », un adjectival **new-o* « de maintenant », « neuf ». Il s'emploiera substantivement pour désigner toute nouveauté, tout nouveau venu. Avec cette dernière valeur, l'utilisation du *-m* d'accusatif va être automatique. Mais dans un état de langue où l'adjectif ne se distingue guère du substantif, où le genre féminin n'existe pas encore, où l'accord en genre n'est encore qu'une redondance maladroite, **new-o*, quel que soit son référent, va représenter une seule et même unité linguistique, et si donc *-m* se manifeste à l'accusatif, il y apparaîtra automatiquement. On aura donc **new-o-m* à l'accusatif, *qu'il s'agisse d'une chose inerte ou d'un être animé*. Si l'entité désignée ainsi se trouve être à l'origine d'une action, la forme **new-o-s* sera naturellement utilisée. Mais si cette « nouveauté » est, par nature, inerte, on n'aura guère d'occasions d'employer à son sujet cette forme en *-s* qui implique une intervention active. Ce qui n'est, au début, qu'une question de fréquence d'emploi finira par se durcir sous la forme d'une opposition entre des unités qui distinguent entre les formes d'accusatif et de nominatif et d'autres unités qui ne le font pas. Parmi les substantifs en *-o*, d'origine adjectivale, les uns vont être susceptibles de réaliser cette distinction : les non-neutres, comme **dom-en-o* « maître », **snus-o* « bru », avec respectivement les nominatifs **dom-en-o-s*, **snus-o-s* et les accusatifs **dom-en-o-m*, **snus-o-m*. Les autres, qui peuvent désormais s'employer comme sujets, à partir du moment où cette fonction n'implique plus automatiquement une activité créatrice, par exemple, avec des verbes d'état, auront une forme indifférenciée pour les deux cas. Cette forme sera naturellement en *-o-m*, la seule antérieurement attestée. Soit, par exemple, l'ancien adjectival **ped-o* dérivé de *ped-* « pied ». Sujet ou objet, il apparaîtra sous la forme **ped-o-m*, celle qui est à la source du skrt *padám* « trace » et du grec *pédon* « sol ».

Ce qui a certainement contribué à la généralisation de l'*m* chez les non-neutres en général et les radicaux neutres en *e/o* est le déséquilibre entre forme et valeur qui résultait de l'emploi d'une forme marquée par *-s* pour le cas du sujet qui s'imposait, non plus seulement comme la forme de l'agent, mais comme la forme hors syntaxe employée dans la nomination et, très vite, dans l'appel, et qui devenait l'accompagnateur obligé de toute forme verbale à un mode personnel. Ceci entraînait pour lui une fréquence bien supérieure à celle du cas objet, dit accusatif, qui était exclu des constructions intransitives. Or cet objet se présentait au départ comme la forme nue du radical, et il était absolument anormal que la forme la plus fréquente fût plus lourde, alors que l'autre continuait à s'identifier formellement avec le radical nu. En d'autres termes, le cas de l'objet continuait à s'identifier avec l'ancien absolutif, alors qu'il avait été privé, au profit de son concurrent en *-s*, de ses prérogatives de nominatif proprement dit et allait l'être bientôt de celles du vocatif. L'extension du *-m* à la catégorie des neutres en *e/o*, la plus susceptible d'expansion, tendait à atténuer ce déséquilibre. L'élimination parallèle du *-s* de nominatif, chaque fois que l'évolution phonique l'autorisait, a été une autre façon de restaurer un état linguistique plus conforme à une loi permanente de l'économie des langues selon laquelle se simplifient les formes fréquentes et, de ce fait, peu informatives.

Il ne faudrait pas croire que la répartition des substantifs entre deux classes de neutres et de non-neutres dériverait d'une façon particulière de percevoir le monde antérieure à l'établissement, dans la langue, d'un comportement différent des unités des deux classes. Si, dès avant l'établissement du genre, on distinguait entre l'eau qui coule $*H_2ek^weH_2$ (lat. *aqua*, danois *å* « rivière ») et l'eau qu'on utilise **welo-den* (grec *húdōr*, danois *vand*, russe *voda*), c'était qu'au-delà d'une identité de substance, on avait affaire à des réalités fonctionnelles différentes. Ces réalités fonctionnelles ont amené **we/oden* à apparaître si exceptionnellement accompagné de la particule d'agent que l'*s* n'y a jamais accédé au statut de désinence. La classe des neutres ne fait que grouper les substantifs qui ont été exposés à la même restriction d'emploi. Les idées que les hommes se font du monde dans lequel ils vivent sont, dans une large mesure, dépendantes des structures linguistiques qu'ils utilisent pour communiquer leur expérience. On ne voit pas qu'à l'inverse, les catégories linguistiques puissent être directement influencées par les croyances, les idéologies ou les modes, toutes sources d'innovations lexicales, mais qui se couleront dans les moules préétablis des classes de monèmes et des schémas syntaxiques pré-existants. Ce n'est pas la pensée qui a créé le langage, mais le langage qui, né des besoins communicatifs les plus divers, a permis à l'homme d'accéder à la pensée.

CHAPITRE XI

LE VOCABULAIRE

Reconstruire le vocabulaire d'une langue disparue pose des problèmes assez différents de ceux qu'on rencontre lorsqu'on cherche à restituer sa grammaire. La grammaire traitant des relations mutuelles des unités linguistiques dans la chaîne parlée, l'interprétation d'un fait grammatical se fera en référence à ce qu'on trouve dans cette chaîne ou, comme on le dit plus savamment, en fonction de ses rapports syntagmatiques. Lorsqu'on se risque au-delà des plus anciens textes, la vérification des hypothèses relatives à la structure grammaticale résulte de la cohérence du modèle reconstruit. Lorsqu'il s'agit du vocabulaire, la compréhension des mots est certes largement assurée par les contextes où on les trouve, mais dès qu'on essaie de remonter au-delà des textes, on devient très dépendant de la nature des réalités désignées, c'est-à-dire des données archéologiques [1].

La parenté.

Il y a toutefois des domaines où le vocabulaire reproduit certaines structures sociales et où, en conséquence, comme en phonologie et en grammaire, l'identification de chaque unité s'y fait par référence aux autres unités du domaine. C'est le cas, en particulier, des relations de parenté. En français, par exemple, *père* implique nécessairement *fils* ou *fille*, *oncle* nécessairement *neveu* ou *nièce*. On peut faire valoir que les rapports constatés entre *père* et *fils* sont dans la nature des choses et que la langue ne fait dans ce cas que refléter la réalité. Mais il n'est pas

1. On renvoie, une fois pour toutes, aux différents dictionnaires étymologiques cités ci-dessus, note 1 du chapitre IV. Parmi eux, il faut mentionner tout particulièrement le *Dictionary of Selected Synonyms* de Carl D. BUCK.

dit que le rapport d'engendrement qui fonde la relation de *père* à *fils* ait été, à une certaine époque et en un certain lieu, celui sur lequel se fondait la cellule de base de la société. Nous avons vu que si la reconstruction de la forme correspondant à *père*, ital. *padre*, angl. *father*, etc., ne posait guère de problèmes, la valeur à donner à cette forme avait dû évoluer considérablement au cours du temps. Cette valeur, nous pouvons la retrouver, ou du moins la deviner, si nous partons des rapports impliqués par les autres termes que nous pouvons reconstruire dans le même domaine. Notre tâche sera, en la matière, facilitée par les progrès réalisés par l'ethnologie, ou, comme on le dit, l'anthropologie sociale, au cours des dernières décennies.

La structure de base de la société est nécessairement celle qui assure sa survivance, celle qui permet la reproduction de l'espèce et le maintien en vie des nouveaux arrivants. La personne la plus directement et évidemment impliquée en la matière est celle qu'on appelle la mère, celle qui porte l'enfant, le met au monde et le nourrit jusqu'au sevrage. Sans doute, pour qu'il y ait enfant a-t-il fallu l'intervention d'un mâle, mais rien ne dit que l'identité de ce mâle soit connue, même de la mère. Dans la mesure où la survie de l'espèce réclame qu'une aide ou une protection soit assurée à la mère et à son enfant par un individu de l'autre sexe, cet individu sera un de ceux dont la parenté biologique avec la mère ne saurait être mise en doute, c'est-à-dire un frère, dont chacun sait qu'il est de la même matrice qu'elle. A défaut d'un frère de la mère, c'est-à-dire de ce que nous désignons comme son oncle maternel, ce pourra être, dans la génération qui précède, le frère de la mère de la mère, donc le grand-oncle maternel de l'enfant. Le seul apparentement valable est donc celui qui passe par la mère. Aujourd'hui encore, en Israël, est considéré comme juif celui dont la mère est juive. On dira que l'apparentement est dans ce cas matrilinéaire.

Tôt ou tard, cependant, le géniteur mâle pourra se manifester, soit qu'il apprécie la présence de la mère de l'enfant, à titre de partenaire sexuelle, soit qu'il voie en elle une source d'avantages divers, qui vont de l'aide domestique aux travaux des champs et à la production d'enfants, susceptibles eux-mêmes d'être des sources de revenu, mais qui résultent également des rapports qu'il établit avec les proches de sa compagne. Même s'il n'échappe à personne qu'il existe un rapport de cause à effet entre l'activité sexuelle et la venue des enfants, le partenaire mâle du couple n'a jamais l'assurance qu'il est effectivement le géniteur d'un enfant déterminé. Sont, au contraire, patents des liens de sang qui unissent l'oncle maternel à l'enfant. Il va donc s'établir entre le partenaire de la femme et le frère de celle-ci un partage des responsabilités dont la nature variera d'une société à une autre. Très fréquemment, c'est l'oncle qui manifeste à l'enfant le plus

d'affection, mais l'on trouve également des situations où c'est lui qui se montre plus sévère que le conjoint.

L'établissement du patriarcat résulte de la décision du partenaire de la femme d'assumer l'entière responsabilité des enfants issus d'elle. Il se veut donc, non seulement le protecteur et l'éducateur de ces enfants, mais également le géniteur. La seule assurance qu'il puisse obtenir à cet égard résultera de la claustration de la femme dans un gynécée, ou un harem, ce dernier impliquant comme on le sait, l'existence de plusieurs partenaires féminines. Lorsque cette claustration se révèle difficile ou économiquement peu désirable, il restera à l'homme le recours de supprimer l'enfant dont la filiation n'est pas sûre en l'exposant à l'appétit des prédateurs. La mythologie grecque nous offre maints exemples d'enfants qui, comme Œdipe, sont, par leur père, abandonnés dans la nature. De ce jour, est reconnu comme l'enfant du chef de famille, celui qu'il a choisi, ce qui ouvre la voie à l'adoption.

Sans affirmer que le passage de l'apparentement matrilinéaire au patriarcat est irréversible, on peut dire qu'il représente un processus évolutif normal, c'est-à-dire qu'il y a quelques chances pour que le mâle, du fait de sa force, cherche finalement à assurer sa domination dans la cellule sociale de base. En ce qui concerne les « Indo-Européens », la reconstruction qu'on peut tenter, à partir des données disponibles, aboutit à un système patriarcal dans lequel subsistent cependant quelques traces d'une structure plus ancienne où s'affirme la priorité des relations à la mère.

Cette cellule de base est ce qu'on appelle la grande famille, groupée sous l'autorité d'un seul, le **pater-*, le protecteur, plus patron que père, normalement le plus ancien de sa génération. Les cadets, lorsqu'ils ne se sont pas éloignés pour fonder leur propre famille, vivent sous sa coupe. Il dispose vraisemblablement de multiples partenaires sexuelles, mais pour la plupart de rang servile, une seule d'entre elles étant habilitée à fournir des prétendants légitimes à sa succession. Si toutefois celle-ci se révèle incapable d'assurer une succession masculine, l'homme pourra avoir recours à une concubine. Le statut de l'épouse résulte sans doute de son appartenance initiale à une autre famille avec laquelle existent, de son fait, des relations privilégiées. Il n'est pas vraisemblable que le mot **māter-* ait été réservé aux relations qui s'établissent entre l'épouse et sa progéniture. Il a dû désigner toute femme du groupe social, esclaves comprises, qui a mis un enfant au monde. Il n'y a donc, au départ, nul parallélisme entre **pater* et **māter*.

Les *frères*, *b^hrātres*, ne sont pas nécessairement nés d'un même géniteur et d'une même mère. Ce sont tous les mâles qui appartiennent à la même génération, aussi bien donc ceux que nous désignerions comme des cousins germains.

Il en va de même des sœurs, _*swe-sōr-es_. Ce dernier terme est, nous l'avons vu, analysable : _*swe_, qui donnera notre pronom réfléchi, n'implique pas autre chose que l'appartenance originelle à la grande famille. Pour les épouses des _*bhrātres_, qui, dans une structure exogamique, viennent nécessairement d'autres familles, le latin présente _uxor_, autre mot en _-sor-_, mais dont on n'identifie pas l'élément initial _uk-_. Tout rapport entre « frères » et « sœurs », même au sens très large qu'ont alors ces termes, serait incestueux. Ce serait, au départ, un péché, sinon contre la morale, du moins contre l'économie du groupe, en ce que les « sœurs » représentent une précieuse monnaie d'échange. Ceci implique naturellement une certaine importance attachée à la virginité des filles.

Même si l'on ne trouve pas trace, à date ancienne, de la mystique chrétienne du pucelage, la coutume germanique de la _Morgengabe_ (vieil-angl. _morgengifu_), don fait par le mari à l'issue de la nuit de noces[2], suggère que la perte de la virginité devait impliquer une sérieuse dévaluation.

Ce n'est que plus tard, dans les langues particulières, là où la grande famille a cédé la place à des organisations plus conformes à nos habitudes contemporaines, qu'on peut relever le remplacement des termes *b^{h}_rāter-_ et _*swe-sōr_, par des formes explicitant l'apparentement physiologique. Le grec a, pour « frère » et « sœur », _adelphós_ et _adelphḗ_, c'est-à-dire deux adjectifs qu'on reconstruit comme _*sm̥-_g^{w}_elb_^{h}_-o_, _*sm̥-_g^{w}_elb_^{h}_-eH_$_2$, c'est-à-dire « d'une même matrice ». Le grec présente effectivement _delphús_ pour « matrice ». De la même racine dérive le nom de plusieurs jeunes animaux, comme angl. _calf_, all. _Kalb_ (pl. _Kälber_ < *g^{w}_olb_^{h}_us-_) « veau ». De façon analogue, mais à date plus récente, l'espagnol a utilisé, pour désigner le frère et la sœur, _hermano, hermana_, du lat. _germanus_, de la racine _*gen-_ qui se réfère à l'engendrement. Dans l'une et l'autre langue, les reflets de *b^{h}_rāter-_ assument la valeur de « membre d'un groupe » social ou religieux.

Dans la grande famille, les mots qu'on reconstruit comme _*sunu-_ et *d^{h}_ug_^{h}_ater-_ et qui survivent, en anglais, pour indiquer la filiation masculine, _son_, ou féminine, _daughter_, marquaient, selon toute vraisemblance, moins une relation spécifique de parent à enfant, que le rapport entre personnes appartenant à des générations successives. Un terme comme le latin _filius_ est, à l'origine, un adjectif, avec un féminin _filia_. On a dû l'ajouter à _*sunu-_ et à *d^{h}_ug_^{h}_əter_ pour préciser qu'il ne s'agissait pas d'un vague rapport d'une classe d'âge à une autre, mais d'une réelle filiation accompagnée d'allaitement. _Fīlius_ est, en effet, dérivé d'une racine *d^{h}_ē(i)_ « téter ». De même le letton a _dēls_

2. Voir Jack Goody, _L'évolution de la famille et du mariage en Europe_, Paris, 1985, appendice II.

« fils » et le russe *ditja* « enfant », *deva* « fille, vierge » de la même racine. Encore aujourd'hui, l'anglais emploie beaucoup plus libéralement le terme *son*, ou son diminutif *sonny*, à l'adresse d'un garçon de la génération suivante que ne le fait le français avec *fils* et *fiston*.

D'un intérêt particulier sont les termes qui désignaient les rapports de la jeune épouse, venue de l'extérieur, avec les différents membres de la grande famille. Dès son arrivée, elle se trouve sous la coupe de l'épouse du chef, ce que nous appellerions sa belle-mère. Elle la désigne comme sa **swe-krū*, où *swe-* désigne naturellement la grande famille. Quant à *-krū*, il ne comporte aucun trait qui le marque comme féminin, qu'il s'agisse du sexe ($-\bar{\imath} < -iH_2$) ou du genre ($-\bar{a} < -eH_2$) et doit, en conséquence, être une forme très ancienne. On la retrouve, aujourd'hui encore, dans all. *Schwieger(-mutter)* « belle-mère ». Le « beau-père », moins directement concerné, était désigné comme le **swe-kuro-*. On soupçonne que le second élément a dû être influencé par l'analogie de **kuro-* « fort, puissant », en avestique *sura* ; on a un *ū* dans le grec *kū́rios* « maître » dont le vocatif *kū́rie* est connu par le *kyrie eleison*, mot-à-mot « Seigneur, aie pitié ! ». **Swekuro-* est à l'origine d'all. *Schwager* qui désigne le beau-frère.

A son arrivée dans la grande famille, la jeune épouse rencontre les « frères », au sens large, de son conjoint, et elle les désigne comme des **daig^w eres* qui subsiste dans le vieil-anglais *tācor*. Le latin a encore *laeuir*, *lēuir* au sens restreint de « beau-frère ». On sait que le *levirat* désigne la coutume hébraïque selon laquelle la veuve devait épouser son beau-frère s'il n'y avait pas encore d'héritier. Les **swesōres* de la famille sont, pour elle, des **$g\partial leH_3$-*, lat. *glōs*, grec *gálōs*, russe *zolovka*. Avec les autres épouses, elles forment l'ensemble des **yenə-tr-*, terme encore attesté en latin sous la forme, modifiée par analogie, de *ianitrīx* ; le grec a *einatéres*, le skrt *yātar-*, le vieux-russe *jatry*.

Un des traits les plus intéressants du vocabulaire indo-européen des noms de parenté est l'archaïsme que représente le terme qu'on traduit généralement au moyen de « grand-père » et qui, sous une forme diminutive, a donné le français *oncle*. Les formes latines sont respectivement *auus* et *auunculus*. On trouve l'équivalent d'*auus* dans le hittite *huhhas* que l'on reconstruit comme **H_2euHo-*, **H_2uHo*.

Il est remarquable qu'*auunculus* ne désigne que l'oncle maternel, l'oncle paternel étant *patruus*. Mais la même restriction n'existe pas pour le simple *auus* dont l'équivalent hittite a toujours spécifiquement la valeur de grand-père paternel ; *auus* lui-même semble s'appliquer de préférence à celui-ci. Il y a là une anomalie qu'on peut, après Emile Benveniste[3], expliquer en rappelant un aspect d'un type d'exogamie. Nous n'entrerons pas ici dans le détail des rapports entre les deux

3. Dans *Le vocabulaire des institutions indo-européennes*, Paris, 1969, livre 2, chap. 3.

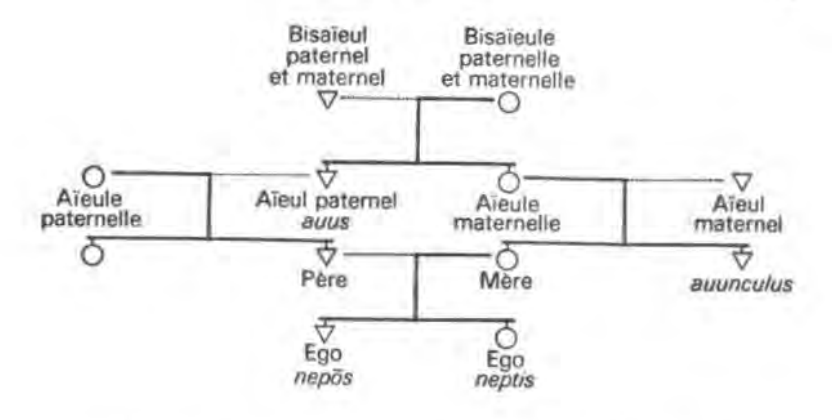


Schéma des relations familiales dans le cas de mariages entre "cousins croisés" combinant, d'aïeul à ego, descendance paternelle et matrilinéarité

Les triangles désignent les hommes, les cercles les femmes. Les rapports matrilinéaires sont marqués par des traits pleins ; les rapports du père aux enfants sont indiqués en pointillé. En italique, on a les formes latines qui indiquent les rapports entre la personne considérée (ego) et ses ancêtres dans des rapports matrilinéaires. On remarquera que le même individu est, tout ensemble, l'aïeul paternel et le grand-oncle maternel, d'où l'ambiguïté du terme *auus* qui le désigne en latin.

moitiés du groupement social à l'intérieur duquel elle fonctionne. Pour résoudre le problème que paraissaient poser les valeurs divergentes d'*auus* et d'*auunculus*, il suffit de rappeler que, selon ce système d'accouplement, un garçon épouse obligatoirement une fille dont la mère est la sœur de son père. En d'autres termes, le mariage se fait toujours entre ce qu'on appelle des « cousins croisés », mais en excluant le cas où le père de la fille est le frère de la mère du garçon. Si maintenant nous considérons une personne qui est le produit de ce mariage, nous constatons que du fait des aléas qui s'attachent à la descendance patrilinéaire, il n'est pas sûr qu'elle soit biologiquement apparentée à son grand-père paternel par l'intermédiaire de son père, mais elle l'est certainement du fait que sa grand-mère maternelle, dont elle descend de toute évidence, était la sœur de ce grand-père. Il y a là une façon élégante de concilier ce qui paraît, à première vue, inconciliable, à savoir le patriarcat d'une part et, d'autre part, l'assurance de l'apparentement biologique qu'assure la descendance matrilinéaire. Définir l'*auus* comme le grand-père paternel est correct du point de vue patriarcal, celui qui prévalait, de toute évidence, dans la société romaine. Mais pour comprendre l'identité formelle partielle

d'*auu-s* et d'*auu-nculu-s* il faut comprendre le premier comme désignant ce que nous concevons comme le grand-oncle maternel ; le second, différencié par l'affixe diminutif *-nculu-*, est l'oncle maternel. Notre rapport grand-oncle–oncle correspond donc, dans les formes latines, au rapport oncle–petit-oncle, la différence s'expliquant du fait de la différence de point de vue : nous avons toutes chances de mieux identifier notre oncle que notre grand-oncle, alors que, dans une société aux structures plus contraignantes et plus hiérarchisées, c'était le grand-oncle qui s'imposait tout d'abord comme le garant d'une descendance biologique assurée.

Ce qui vient d'être dit de l'oncle et du grand-oncle vaut, *mutatis mutandis*, pour le neveu et le petit-neveu. On doit partir d'une forme **népot-*, avec pour la nièce, une variante **nept-* suivie du suffixe de sexe féminin -iH_2-, d'où en sanskrit *napāt* et *naptīs* et, en latin, *nepōs, neptis*. Dans les sociétés indo-européennes les plus conservatrices, il s'agit toujours des enfants de la sœur ou des descendants matrilinéaires en général. Les relations marquées par *auus* et *auuculus* sont donc, dans l'autre sens, exprimées par *nepōs*. Comme dans le cas d'*auus*, les termes *nepōs* et *neptis* vont, dans le cadre du patriarcat, pouvoir désigner aussi bien les produits d'une descendance masculine, mais, en latin, certains emplois rappellent le traitement bienveillant que l'oncle accordait aux enfants de sa race face à la stricte discipline imposée par le *pater familias*. Le *nepōs* est un peu l'enfant gâté, et, encore aujourd'hui, nous désignons comme le népotisme, le comportement de celui qui favorise systématiquement les membres de sa famille.

La grande famille, dont on vient d'esquisser à grands traits le fonctionnement, avec sa structure strictement hiérarchisée, a joué un grand rôle dans l'expansion indo-européenne. Celui qui est à sa tête, le **pətēr*, est également désigné comme le chef de la maison, en grec *despótēs*, interprété comme formé de *dems-*, génitif de *dem-* « maison », et de *pot-* qui désigne celui qui a le pouvoir : le latin *potest* « il peut » s'analyse clairement comme *pot-* « puissant » plus le verbe « être ». C'est naturellement de *despótēs* que dérive notre *despote*, et ceci explique la tentation que pouvait avoir tout cadet d'aller fonder ailleurs sa propre famille pour échapper à la domination du chef. Le terme *domination* lui-même se réfère au comportement du *dominus*, chef de la maison, *domu-s*. C'est bien là le type d'organisation sociale que révèlent les kourganes et que les conquêtes indo-européennes ont dû imposer à de larges portions de l'Eurasie avant que l'évolution des économies et l'influence des substrats des divers peuples subjugués déterminent un relâchement des disciplines patriarcales.

L'avènement du christianisme, avec l'accent mis sur le salut individuel, a abouti à affaiblir la puissance des liens familiaux. La

possibilité de tester a permis à l'individu de priver ses descendants de leur héritage au profit des institutions monacales dans l'espoir d'un meilleur sort dans la vie de l'au-delà[4].

On ne peut guère reconstruire de terme qui désigne la grande famille. En latin, on dispose de *gens*, de la racine *gen-* « engendrer », pour désigner les gens identifiés comme de même origine masculine. Mais, à Rome, la *gens* n'est plus la cellule de base. Celle-ci est la *familia*, c'est-à-dire la maisonnée, caractérisée par la présence d'esclaves, les *famuli*. Ceci veut dire que, même lorsqu'ils demeuraient à proximité relative de l'aîné, dans la même cité, les cadets ont su relâcher les liens de soumission qui l'enchaînaient à lui.

La société.

De façon générale, la nature et la structure des cellules sociales ont dû évoluer au cours du temps. L'unité immédiatement supérieure à la grande famille a été désignée à partir d'une racine **weik-*. Elle peut se présenter sous la forme **wik-*, skrt *viç-* « établissement », devenue, par un processus analogique fréquent (l'accusatif **wikm̥* passant à **wikim*), **wiki* en balte et en slave, d'où le russe *veś* « village ». On la trouve aussi dans un dérivé en *-e/os-*, got. *weihs* « village », et un autre en *-e/o-*, lat. *uīcu-s* « pâté de maison », d'où le dérivé *uīcīnus* « voisin », skrt. *veçā-s* et gr. *(w)oîkos* « maison ». Le mot se réfère évidemment à l'ensemble matériel des bâtiments, mais sa valeur sociale se reflète dans skrt. *viçpatis* « chef de village » et lit. *vẽšpats* « seigneur ».

En germanique, on trouve un mot qui vaut pour le village comme pour la résidence individuelle, c'est l'anglais *home*, correspondant au got. *haims*, au lit. *kaimas*, au grec *kṓmē*, tous « village » ; le russe *sémja*, qui en dérive, désigne la famille. La racine **kei-* semble être celle qui veut dire « s'étendre pour dormir », ce qui est effectivement le sens du grec *keî-mai*. Avec une autre dérivation, on a got. *heiwa-* dans *heiwafrauja* « chef de famille », vieux-h.-all. *hīwo* « mari » ; on y rattache le lat. *cīui-s* « citoyen » où la valeur politique est probablement tardive. L'ensemble des citoyens forme la *cīuitās*, la cité. Ce mot a déplacé l'ancien terme occidental désignant le peuple, en osque *touto*, en irl. *túath*, en got. *þiuda*, à partir duquel une dérivation adjectivale a donné, en latin carolingien, *thiudisce* « populaire » qui s'opposait à *romanice* « de langue romane ». De ce mot dérive, on l'a vu, all. *deutsch* « allemand » et, en fr. *tudesque*. Un traitement dialectal d'un latin **teuta* a pu aboutir à *tōta*, d'où un masculin *tōtus* avec le sens de « tout entier ». A noter que le russe *veś* « village » a également le sens de « tout ».

4. C'est l'essentiel de la thèse de Jack GOODY, dans *L'évolution...*

On aura constaté la facilité avec laquelle on passe de valeurs strictement matérielles à des valeurs abstraites avec implications sociologiques, voire politiques.

On peut penser qu'à une certaine époque, l'unité matérielle correspondant à **wik-* a pu coïncider avec le clan. Au-dessus du clan et avant qu'on n'en arrive à la cité, il faut placer la tribu. Mais il n'est pas possible de reconstruire, pour cette notion, un mot appartenant à plus d'un groupe. Le latin *tribus* où l'on a voulu interpréter *tri-* comme faisant, de la tribu, le tiers du peuple, n'est pas attesté hors de l'italique. Un terme qui jouit d'une diffusion un peu plus large est celui qui correspond à l'all. *Heer* « armée ». En germanique, il a un sens strictement militaire, le verbe correspondant ayant la valeur de « piller », « dévaster ». Il en va de même en balte, où il est peut-être emprunté au germanique. En celtique, le sens des mots correspondants est plutôt « troupe », « communauté » et le nom du peuple gaulois des *Petru-corii*, qui, nous l'avons vu, a donné *Périgord*, évoque, plutôt qu'une armée proprement dite, une tribu qui se déplace avec armes et bagages. La même valeur incertaine caractérise les antécédents d'angl. *folk*, all. *Volk* « peuple », mais, anciennement, une troupe guerrière, comme on le voit par le slave, où le mot, qui est en russe *polk*, très tôt emprunté aux Germains, désigne l'armée.

Ni la cité au sens politique du terme, ni la ville au sens matériel, ne figuraient dans la vision du monde de l'indo-européen commun. Les termes qui correspondent aujourd'hui à ces notions dérivent de désignations vagues de lieu, comme l'all. *Stadt*, d'habitation, comme le scandinave *by*, d'enclos, comme l'anglais *town* ou le russe *gorod*, de ferme, comme le français *ville*. Seul le grec *pólis* semble pouvoir être rapproché d'autre chose, notamment le skrt *pūr-*, *pur-*, qu'on retrouve à la finale de *Singapour* et dans de nombreux noms de ville de l'Inde. Le lit. *pilis*, avec le sens de citadelle, indique la valeur primitive du terme.

Sur la structure des sociétés indo-européennes anciennes, on dispose aujourd'hui des résultats de recherches poursuivies notamment, en France, par Georges Dumézil[5]. Par un rapprochement des données recueillies à deux extrémités du monde indo-européen, à Rome et dans l'Inde ancienne, il a dégagé l'existence de trois classes, l'une détentrice du pouvoir religieux, celle des prêtres, la seconde, celle des guerriers, et la troisième chargée de ce qu'on appellerait l' « intendance », c'est-à-dire la satisfaction des besoins matériels, donc soit les agriculteurs, aux époques relativement récentes, soit, plus anciennement, les pasteurs. L'ordre dans lequel on vient de les citer marque une hiérarchie.

5. Voir, notamment, *Les dieux souverains des Indo-Européens*, Paris, 1977.

Les dieux.

Une hiérarchie parallèle existe entre les divinités qui sont censées présider aux activités des trois classes. Rome d'une part, l'Inde védique de l'autre, sont parmi les établissements indo-européens, ceux où sont le mieux représentés les corps de prêtres. On a même tenté de rapprocher, sinon d'identifier, les flamines romains et les brahmanes indiens, en dépit de difficultés formelles. Ce sont également deux des branches où est bien établie la forme qui a donné en latin *rēx*, d'où le français *roi*, en skrt *rāj-*. Il faudrait sans doute adjoindre au groupe les Celtes, qui connaissent le corps des druides et ont conservé, sous la forme *-rīx*, la désignation du « roi ». Toutefois le parallélisme structural ne s'étend ni aux formes ni aux détails de l'organisation. Le sommet de la hiérarchie est occupé, à Rome, par *Jupiter*, où le radical permanent se reconstruit comme **dye/ow-* (le génitif est *iou-is ; -piter* valant bien entendu *-pater-*, « le protecteur ») ; dans l'Inde védique, on trouve là une dualité *Mitra-Varuṇa* qui combine en les opposant le compromis (*Mitra*) et la violence (*Varuṇa*), c'est-à-dire les deux façons de régler les différends. Ensuite vient le dieu de la guerre, *Mars* d'une part, *Indra* de l'autre, et finalement, correspondant à la troisième classe, des divinités moins bien définies comme, à Rome, *Quirinus*.

On pourrait être tenté de faire valoir que cette tripartition sociale se rencontre, en fait, un peu partout, avec, en tête, les détenteurs des pouvoirs magiques, ensuite ceux qui représentent la force brutale, et enfin ceux qui ne s'imposent que parce qu'on dépend d'eux pour la subsistance. Elle est, par exemple, celle qu'on retrouve avec les trois Etats de l'ancien régime en France, et généralement dans l'Europe médiévale. Elle ne prend de sens, pour la société indo-européenne, que lorsqu'on la rapproche de ce qu'on pourrait vouloir considérer comme des déviations par rapport à elle, celles, par exemple, qu'on constate chez les Germains, où le dieu suprême, *Odin* (*Wotan* sous sa forme allemande), est celui des guerriers qu'il reçoit, après leur mort héroïque, dans son Walhalla. Dans ce panthéon, *Týr* (vieil-angl. *Tīw*), l'équivalent formel du Jupiter latin et du *Zeus* grec (tous de **dye/ow-*), fait pâle figure après avoir cédé sa majesté à *Odin* et ses foudres à *Thor*. Dans l'adaptation des noms de jours de la semaine latine, le *mardi*, jour de *Mars* est, curieusement, devenu le jour de *Tīw* (angl. *Tuesday*), le *mercredi*, jour de *Mercure*, celui d'*Odin-Wotan* (angl. *Wednesday*, dan. *onsdag*), le *jeudi*, jour de *Jupiter*, celui de *Thor* (angl. *Thursday*, dan. *torsdag*).

Il se peut aussi que l'opposition entre la classe des guerriers et celle des pasteurs-agriculteurs s'atténue, dans la mesure où tout homme libre peut être considéré comme un combattant en puissance, le fer

servant concurremment pour l'épée et pour la charrue (lat. *ense et aratro*).

On peut donc retenir la tripartition en prêtres, guerriers et paysans comme un schème social hérité, par certaines des différentes branches, de la période de l'indo-européen commun. Toutefois la comparaison ne permet guère de préciser ce que pouvait être, avant le départ des Anatoliens, le panthéon indo-européen. A l'exception de Jupiter, dont le nom, sous diverses variantes, se retrouve largement diffusé, les désignations des dieux, sinon leurs fonctions, ne correspondent plus d'une branche à l'autre. On peut donc poser un dieu du ciel diurne et de la foudre **dye/ow,* évidemment accompagné d'autres divinités, mais dont on ne saurait reconstruire la désignation, voire la fonction exacte. On peut, dans l'optique dite païenne, poser un dieu chaque fois que vous échappe un chaînon de la causalité et qu'on suppose une force agissante. Quand on se rappelle ce qu'est en fait un nom de genre neutre, on peut estimer que tout non-neutre est un candidat potentiel à la divinité. Le feu, par exemple, dont on ne voit pas pourquoi il jaillit tout d'un coup de matériaux où il couve, est un excellent candidat, et, effectivement, *Agni* « le feu » est un des grands dieux indiens. Pour la notion même de « dieu », on a des représentants de **deiwo-* — variante de **dye/ow,* qui désigne le dieu du ciel et du jour — en latin (d'où le français *dieu*), en celtique, en balte et en sanskrit. Mais le germanique, avec got. *guþ*, angl. *god*, all. *Gott,* a innové ; de même le slave, avec *bogŭ,* russe *bog*, emprunté sans doute à l'iranien, et le grec, avec *theó-s*.

Ce qui est bien attesté est l'opposition entre l'homme (l'espèce), qui est « le terrestre », et le dieu, qui est céleste. Lat. *homo,* d'où *homme,* dérive de la racine qui désigne la terre, représentée, par ailleurs, en latin, par *humus*. Les formes celtiques, du type de l'irlandais *duine,* sont exactement de même origine, la base commençant, comme on l'a vu, par un phonème complexe combinant une articulation dorsale de type g^h, d'où le *h* du latin *homo* et le *g* de l'équivalent vieil-anglais *guma,* et une articulation apicale d^h qui a donné le *d-* celtique. On a, bien sûr, en face, **dye/ow-* qui est, tout ensemble, le ciel et la divinité suprême. La même opposition peut se manifester par l'utilisation, pour désigner l'homme, de « mortel », comme le persan et l'arménien *mard,* le grec *brotós,* pour **mrotós*.

Les déplacements.

La comparaison renseigne mal sur la classe des guerriers. Les techniques martiales varient considérablement au cours des siècles, ce qui entraîne des renouvellements du vocabulaire. Toutefois, le cheval et le char qu'il traînait, qui ont été les grands instruments de conquête

des peuples indo-européens, sont largement attestés sous des formes qu'on peut rapprocher les unes des autres. La désignation la plus ancienne pour le cheval est probablement celle qui survit dans l'angl. *mare* « jument » et qui est bien représentée en germanique et en celtique. C'est elle qu'on a dans *maréchal*, au départ le serviteur *(skalk-)* préposé à la garde et au soin des chevaux. On la trouve en Extrême-Orient, en chinois, par exemple, sous la forme *ma*. Mais le mot le plus répandu est une forme de type adjectival **ekwo-*, lat. *equos*, grec *híppos*, qu'on retrouve dans *équestre* et *hippique*. Le rapprochement, sémantiquement tentant, avec gr. *ōkú-s*, skrt. *āçú-s* « rapide », fait formellement difficulté. L'anglais *horse*, qui représente la forme la plus courante en ancien germanique, est rapprochée du latin *cursus* « course ». Ce mot remonte à une racine à laquelle on rattache les désignations du char en latin, *curru-s*, et en celtique, *carro-s*. D'une autre racine **ret*, du sens de « courir », « rouler », on dérive le latin *rota* « roue » (**rot-eH₂*), l'all. *Rad*, lit. *ratas* de même sens. Nous avons rencontré ci-dessus la désignation sanskrite du char au moyen d'un adjectival tiré de **rot-eH₂-*.

La notion de se déplacer ou de transporter en char s'exprimait au moyen de mots dérivés d'une racine **wegʰ-* représentée par skrt *vah-a-ti* « il transporte en char », lat. *ueh-o* « je transporte », vieux-slave *vezǫ*, même sens. En grec, la chute du *w-* initial dans l'équivalent *(w)ekhō* a amené une homonymie avec un autre verbe *ekh-ō*, du sens de « tenir », provenant d'une racine **segʰ-*, ce qui a limité son extension. La racine est attestée également dans les désignations de véhicules. Avec un suffixe *-e/os-* on a, en grec *ókhos-* « char » et, avec *-no-*, des formes du sens de « voiture », en irlandais, *fén*, et dans les langues germaniques, avec suéd. *vagn*, all. *Wagen*, anglais *wain*, *wagon*, tous de **we/ogʰ-no-*. On trouve aussi cette racine dans des mots du sens de « chemin », comme got. *wigs*, all. *Weg*, anglais *way* et le latin *uia*, d'où le français *voie*.

Le déplacement sur l'eau, certainement beaucoup moins fréquent, a pu être désigné au moyen de termes ayant la valeur de flotter, comme on le constate dans le cas du grec *pleō* (< **plewō*). La racine **pleu-* est, en outre, attestée en slave, russe *ply-t*, *plav-at'-* avec les sens de « flotter », « nager » et « naviguer ». En germanique elle a pris la forme **pleu-d-*, d'où angl. *float*, all. *fliessen*, suédois *flyta*. En balte, on a avec les mêmes valeurs, des formes dérivées de **pleu-k-*. L'embarcation elle-même est désignée au moyen d'une forme dérivée d'une racine **neH₂w-* qui apparaît dans le skrt *naus*, le grec *naû-s*, l'irl. *nau*, le vieil-isl. *nó-r*, l'arménien *naw*, le latin *nāu-i-s*. Il existe en germanique une forme en *-k-* : vieil-anglais *naca*, all. *Nachen* « nacelle, barque », qui pourrait s'expliquer par le durcissement de H_2 devant le *-s* de nominatif, *-H₂w-* ayant été traité comme H_3, c'est-à-dire [χ^w]. La

trace du *-w-*, qu'on retrouve dans le vieil islandais *nǫkkve*, serait due dans ce cas à une extension analogique à partir des autres cas. Le mot a été ultérieurement muni d'un suffixe *-on-*, probablement de valeur diminutive, encore attesté par le *-en* du mot allemand.

L'élevage.

La classe des pasteurs-agriculteurs, en tant que telle, n'a pas laissé de traces linguistiques aussi claires de son importance à l'époque de l'indo-européen ancien. Mais une portion considérable du vocabulaire qu'on peut reconstruire se réfère à l'élevage et à l'agriculture. Il est clair que bovins et ovins ont joué un rôle considérable qui suppose leur domestication.

Les troupeaux représentaient l'essentiel du patrimoine et leur protection contre les razzias, fréquentes chez un peuple qu'on peut supposer fort peu policé, a été à la source d'un vocabulaire qui s'est finalement étendu aux domaines les plus divers. Si les soins quotidiens du bétail étaient, de toute évidence, l'affaire de la plus humble des trois classes de la société, sa défense contre l'agression ne pouvait manquer d'impliquer l'intervention des guerriers.

Pour l'essentiel, ce vocabulaire dérive de deux racines distinctes[6]. Nous avons déjà rencontré la première, *peH_2, lorsque nous avons tenté de préciser le sens primitif du mot **pətēr* interprété, non comme le géniteur, mais comme le protecteur du clan. Avec le sens spécifique de protecteur, on trouve un doublet avec voyelle pleine du radical, par exemple, dans skrt *narām ... pātá* « protecteur des hommes » ou dans le composé de même sens *nr̥-pātŕ̥-* (où *nr̥-* désigne les hommes des deux classes supérieures). Avec un suffixe d'instrument, on a, dans la même langue, *pātra-* « réservoir », et, en germanique, le gotique *fodr* « fourreau ». Le mot français *fourreau* est dérivé de la base germanique. Pour désigner, non plus ce qui protège, mais ce qui est protégé, avec spécialisation dans le sens de la nourriture et divers suffixes, on a lat. *pānis* « pain », angl. *food* « nourriture », et, pour le fourrage, lat. *pābulum*, all. *Futter*, angl. *fodder*, avec, ici encore, le français *fourrage* dérivé de la base germanique. L'anglais *foster*, au départ « nourrir », puis « élever », est utilisé comme préfixe pour indiquer l'apparentement nourricier. Avec une suffixation en *-s-* ou en *-sk-*, on trouve en slave et en latin les mots qui ont rapport au pacage : pour « paître », dans les deux sens du terme, « mener à la pâture » et « brouter

6. On trouvera un exposé détaillé de la question dans un long article de Françoise BADER, De « protéger » à « razzier » au néolithique indo-européen : phraséologie, étymologie, civilisation, *Bulletin de la Société de linguistique* 73, 1978, p. 103-219.

l'herbe », le latin a *pasco(r)* d'où le français *paître*. Celui qui paît le bétail est, en latin, *pastor*, d'où le français *pâtre*.

Il existe une racine parallèle, avec H_3 au lieu de H_2 et adjonction d'un élément *-y-* qui apparaît dans skrt *nṛ́-pīti-* à rapprocher de l'équivalent *nṛ-pātí*, ci-dessus. De la même base, avec diverses suffixations, on a, en grec, *pōü* « pâturage » et *poimḗn* « berger ». De la notion de « protéger », on est passé à celles de « préserver », « observer (ses obligations) », d'où le latin *pius* « pieux » (< **pī-us*) et un verbe *piāre* qui implique, par l'observance d'un sacrifice, la recherche de la protection d'un dieu ; *expiāre* « expier », c'est aller jusqu'au bout dans le sacrifice.

L'autre racine qui recouvre la notion de protection et des valeurs connexes apparaît sous trois formes différentes : **swer-*, **wer-* et **ser-*, ce qui rappelle les variations de l'initiale de *six* (ci-dessus, p. 205) et celles de *swe-/se* « soi-même » employé comme préfixe pour désigner la grande famille (voir, ci-dessus, p. 233). Les trois variantes se partagent, avec **peH₂*, différemment selon les langues, un champ sémantique assez vaste. La base **swer-* est présente dans le grec *horáō* « voir », au départ, sans doute, « garder », « regarder », qui apparaît dans *panorama* (**pan-(h)ora-ma*), étymologiquement « tout ce qu'on voit ». La base **wer-* est présente en indo-iranien, en hittite et, surtout, en grec et en germanique où **peH₂*, au sens propre de « protéger », a été éliminé à son profit. On la retrouve dans all. *wehren* et dans le premier élément de *Wehrmacht*, littéralement « puissance de protection », c'est-à-dire « armée ». Le latin a *uereor* « vénérer » avec une évolution sémantique analogue à celle qu'on a relevée ci-dessus à propos de *pius*. Avec un suffixe *-dʰ-*, on a des formes germaniques représentées dans l'anglais *ward* « gardien », « protection », *ward off* « écarter (un danger) » et attestées très largement dans les langues romanes, avec durcissement régulier de *w-* en *g-*, dans *garde*, *garder*, *regarder*, etc.

La forme **ser-* apparaît avec divers suffixes. Avec vocalisme zéro de la racine, **sṛ-gʰ-* a donné all. *Sorge* « soin, souci » et angl. *sorrow* qui n'a plus que le sens de « souci ». Le suédois a *sorg* dans un sens, *omsorg* dans l'autre. La même forme, en grec, a assumé deux valeurs distinctes, celle de « commander » dans *árkhō*, d'où le participe présent dans le nom des *archontes*, hauts magistrats athéniens, puis celle de « commencer » probablement par l'intermédiaire de « prendre l'initiative ». Noter qu'avec un vocalisme *-o-*, on a gardé, dans *órkhamos*, le sens de « gardien ». Avec un suffixe *-w-*, on a le latin *seruāre* « conserver ». Quant à *seruos* « esclave », on l'a expliqué de façons diverses, notamment comme le prisonnier « préservé (de la mort) ».

Le nom de la vache **gʷeH₃-* se retrouve partout, de l'irlandais *bó* au

sanskrit *-gu-s*, *gávi* (locatif) avec le traitement fréquent du $*g^w$ en *b*, comme en latin (*bōs*), en grec (*boûs*) et en *k* en germanique (anglais *cow*, all. *Kuh*) et en arménien (*kov*). Le latin *uacca*, d'où le fr. *vache*, présente le traitement régulier de $*g^w$ initial ; le *c* géminé est la marque d'emplois dits « expressifs » du type de ceux auxquels reste soumis ce même mot dans l'usage contemporain (par exemple, *vachement* comme marque de superlatif) ; il pourrait résulter d'un durcissement de la « laryngale » H_3 devant l' *-s* de nominatif ; le *-a-* correspond à ce qu'on noterait *ə, c'est-à-dire le degré réduit de la voyelle en contact avec la « laryngale ». On posera donc $*g^weH_3\text{-}s > *g^wok\text{-}s$; mais, avec le degré réduit de la voyelle lorsque l'accent va porter sur le suffixe de genre féminin, on aura $g^{wo}k(k)\acute{e}H_2$, d'où, régulièrement, *uacca*.

Il ne peut faire de doute que l'élevage bovin a joué un rôle considérable dans l'économie de l'ancienne société indo-européenne, mais surtout peut-être comme source de peaux, sinon très vite de cuir, et de viande. « Cuir », en grec, se dit notamment *skûtos*, d'une racine qui veut dire « couvrir » : skrt. *sku-nâ-ti* « couvrir », grec *skûlon* « dépouille », avec des mots qui désignent l'objet ou le bâtiment qui cache ou protège comme le bouclier, vieil-isl. *skaunn* et peut-être latin *scūtum*, ou le hangar, vieux-haut-all. *scūr* « hangar », ou la grange, dans la même langue, *scūra* (ou *sciura*, d'où l'all. *Scheuer*). On en rapproche le *-scūr-* du latin *obscūrus* « obscur ».

On est moins sûr que le lait ait joué un rôle aussi considérable qu'aujourd' hui. Pour désigner ce liquide et la traite qui le fournit, il devait y avoir deux radicaux en compétition, l'un **(g)lak-t-* qui a donné grec *gala* (gen. *galaktos*), lat. *lac* (génitif *lact-is*), l'autre **me/olg* représenté par got. *miluk-s*, angl. *milk*, all. *Milch*, latin *mulgeo* « je trais », vieil-anglais *melcan*, alternant avec *me/olk-* attesté dans russe *moloko* « lait » et lat. *mulceo* « je palpe » (sans doute doublet de *mulgeo*). On ne trouve pas, dans les différentes langues, à date ancienne, de termes qui permettent de distinguer le beurre des autres graisses animales : en espagnol *mantequilla* « beurre » est un dérivé de *manteca* « lard » ; le mot germanique ancien, conservé en Scandinavie, dan. *smør*, se rattache à une racine qui veut dire « oindre ». En ce qui concerne le fromage, il ne semble pas qu'on puisse reconstruire une forme ancienne ; les emprunts ultérieurs, de langue à langue, semblent indiquer que les techniques de production du fromage étaient peu développées. Les formes celtiques et germaniques occidentales (all. *Käse*, angl. *cheese*) dérivent du latin *cāseus* qu'on rapproche, de façon hasardée, du russe *kvas*, la boisson alcoolique connue, *kvapusta* « choucroute », mots impliquant une fermentation. Le grec *tūrós* a été rapproché d'une forme avestique désignant le « petit lait ». On est toutefois tenté par un apparentement avec le russe *syr* (< *sūro-s*), vieux-prussien *suris*, de la famille de l'angl. *sour*, all. *sauer* et français

sur venu du germanique. Il faudrait, dans ce cas, supposer un emprunt fait par le grec à quelque langue intermédiaire justifiant la correspondance parfaitement inattendue de *t-* à *s-*. Le mot *beurre*, tout comme ses équivalents angl. *butter* et all. *Butter*, provient, en dernière analyse, par l'intermédiaire du latin, du grec *boútūros*, littéralement « fromage de vache ».

Nous avons déjà rencontré le nom du mouton **H*$_3$*ewi-*, qui s'est moins bien conservé que celui de la vache, mais qu'on retrouve en irlandais, en roumain, et en anglais pour la femelle sous la forme *ewe*, mais surtout en slave : russe *ovca*. Le français a *ouailles* pour les agneaux du seigneur.

Pour l'agneau, la forme représentée par le latin *agnus* a une certaine extension, avec le grec *amnós*, l'irlandais *úan*, le vieux-slave *agnĭcĭ*, le vieil-anglais *ēanian* « mettre bas » en parlant de la brebis ; mais il n'est pas facile de poser une forme originelle unique : le latin et le grec supposent **agno-* ou *^{o}*gno-*, l'irlandais **ogno-*, le slave **ōgno-* ou **āgno-*, le vieil-anglais *ēanian* « agneler » **ag*^{hw}*no-* ou **ak*^{w}*no* (ou *^{o}*g*^{hw}*no-*, *^{o}*k*^{w}*no-*) qui moyennant certaines dissimilations pourraient remonter à **H*$_3$*g*^{hw}*no* [χ^{w}...g^{w}no] ou **H*$_3$*g*^{hw}*no* avec différents vocalismes.

Le petit bétail, composé surtout d'ovins, était désigné au moyen de **peku*, lat. *pecu* et ses dérivés, d'où ital. *pecora* « mouton », dan. *får* « mouton », d'une racine qui désigne la toison, la laine (gr. *pékos*) qu'on arrachait du dos du mouton au lieu de le tondre[7]. Cette racine se retrouve dans grec *pékō* « je peigne », lat. *pecten* « peigne », *pectus* « la poitrine », c'est-à-dire la partie velue. Mais **peku* a vite pris le sens plus général de « bétail » et, comme le bétail était un des éléments essentiels du troc, la valeur de « monnaie d'échange » et plus tard d' « argent », d'où nos formes contemporaines *pécuniaire*, *impécunieux*, etc. Le got. *faihu* n'a que la valeur d' « argent, moyen de paiement ». All. *Vieh* désigne le gros bétail, mais angl. *fee* n'a plus que le sens de « rémunération ». Il n'y a pas de raison valable pour poser, avec Emile Benveniste[8], la priorité du sens de « moyen de paiement ».

On a pu se demander si le porc était déjà domestiqué à l'époque de la communauté, parce que le terme **porko-* qui désignait le jeune porc domestique ne connaît, en indo-iranien, que des attestations douteuses. En tout cas **suH-* a été longtemps général pour désigner l'espèce, avec des produits en *sū-*, en *suk*, analogiques d'un traitement devant *-s*, et en *su-*, analogiques de celui de **suH-* devant voyelle. En germanique et en slave, c'est un dérivé en *-īn* qui s'est établi : all.

7. Cf. Jean HAUDRY, Linguistique et vocabulaire indo-européen, dans *E. Benveniste aujourd'hui, Actes du Colloque international de Tours*, 28-30 septembre 1983, tome II, p. 77-84 ; voir p. 82.

8. Dans *Le vocabulaire...*, livre 1, chap. 4.

schwein, russe *svin'ja*, sauf, en germanique, pour la femelle qui est angl. *sow*, all. *Sau*. **Porko-*, sous sa forme simple ou celle de dérivés, a une belle extension et continue souvent à désigner l'animal jeune ou même, comme dans l'anglais *farrow*, la portée.

La chèvre est désignée de façon assez variée dans les différentes langues. On peut rapprocher le grec *aíx* (génitif *aigós*) et l'arménien *aic*. Le latin *haedus*, qui désigne le chevreau, s'identifie avec le mot germanique qui désigne l'espèce : *g^haido-*, got. *gaits*, angl. *goat*, all. *Geiss*. Le latin *caper* (et le féminin *capra* d'où *chèvre*) s'identifie formellement avec le grec *kápros* qui s'applique au sanglier. Il n'est pas rare que les mots qui désignent les mâles se retrouvent d'une espèce à une autre. Le celtique a une forme **gabor-* qu'on pourrait rapprocher du latin *caper* si l'on peut penser que le mot a été emprunté à une langue occidentale pré-indo-européenne où [k] et [g], [p] et [b] n'étaient pas distincts. On a relevé, en particulier dans des noms de lieu, des formes désignant des rochers ou des pierres où *ka(n)t(a)* alterne avec *ga(n)d(a)*. Parmi eux, se trouvent les toponymes et anthroponymes fréquents du type *Chantemerle*, *Chanteraine*, où le second élément a dû être modifié sur l'analogie de noms d'animaux comme *merle*, ou *raine* (c'est-à-dire « grenouille »), de façon à l'apparier avec le premier interprété comme une forme de *chanter*. Dans l'expression argotique *ramasser un gadin* « trébucher et tomber à terre », *gadin* est connu, dans certains usages provinciaux, avec le sens de « pierre ». Toutes les désignations de la chèvre que nous avons énumérées ci-dessus présentent une voyelle de timbre *a* permanent qui, en indo-européen, est l'indication d'une création assez tardive ou d'un emprunt à quelque autre famille linguistique. Doit-on interpréter tout cela comme la marque d'une domestication tardive de l'animal ? Le français *bouc* doit être rapproché de formes celtiques et germaniques du type **bokk-*, **bukk-*, avec un *k* géminé, et un vocalisme fantaisiste qui marque la forme comme familière et expressive. Qu'on pense, en français, à *bique*, *biquet*, *biquette*, avec l'alternance, fréquente, dans ce niveau de langue, entre le [i], diminutif et plus féminin, en face de [u] ou [o] plus mâle ou plus brutal, dans des paires comme *bigre* et *bougre*, *fiche* et *foutre*, *chiper* et *choper*. On notera toutefois le *e* de l'ital. *becco* « bouc ».

L'âne, désigné un peu partout en Europe au moyen d'un terme dérivé des mots latins *asinus*, *asellus*, ne figurait certainement pas parmi les espèces domestiques indo-européennes. Il en va de même, bien entendu, de l'hybride, le mulet, dont le nom dérive assez généralement du latin *mūlus*.

Parmi les animaux de basse-cour, la poule et le coq présentent des désignations qui ne permettent pas de reconstruire des formes communes. Il y a des termes de valeur diminutive ayant désigné les

jeunes animaux, généralement préférés sur les marchés, *poule*, *poulet*, par exemple, qu'on rapproche de *poulain*, en anglais *foal*, et de *pouliche*, en anglais, *filly*, et aussi du lat. *puer* « enfant » ou du grec moderne *-poulos*, qui veut dire « fils » dans les patronymes. D'autres formes, comme *coq*, anglais *cock*, les formes *kokos*, *kokosz* du serbo-croate et du polonais et les dérivés anglais *chicken*, suédois *kyckling* pour le poulet, rappellent les sons produits par l'animal. Quant aux formes germaniques, all. *Hahn* « coq », angl. *hen*, dan. *høne* « poule », elles se rattachent à la racine attestée dans le latin *cano* « je chante ».

Pour l'oie et le canard, au contraire, on rencontre des formes qui nous permettent de remonter dans le temps. Bien entendu, on ne peut guère savoir si les espèces en cause étaient ou non déjà domestiquées. Pour l'oie, on peut reconstruire *g^hans-* attesté en sanskrit comme *haṁsa-*, dans le grec *khḗn*, le latin *anser*, avec chute de l'*h* dans ce terme rural, l'allemand *Gans*, l'angl. *goose*, lit. *žąsis*, le vieux-slave *gąsĭ*, le russe *gus*. Le mot sanskrit vaut aussi pour le cygne, et c'est cet animal que désigne l'irlandais *géis*. Pour le canard, on posera *H_2enH_2ti-*, *H_2neH_2ti-*, *$H_2\mathring{n}H_2ti$-* avec trois variantes accentuelles. A la première, on rattache lat. *anas* (gen. *anatis*), vieux-fr. *ane*, all. *Ente*, dan. *and*, lit. *ąntis*, vieux-slave *ąty* ; à la seconde gr. *nẽtta*, de **nātya* ; à la troisième, skrt. *ātís* qui désigne un oiseau aquatique. Le français *cane*, d'où le dérivé *canard*, représente l'ancien *ane* affecté par *cancan*, imitatif des sons produits par l'animal. L'anglais *duck* est « celle qui plonge » et le masculin correspondant, *drake*, doit remonter à un emploi humoristique du mot qui désigne le dragon, en vieil-anglais *draca*.

La faune.

L'archéologie nous révèle que le chien a été très tôt domestiqué. Ceci se reflète dans la possibilité de reconstruire une forme ancienne dont sont dérivés plusieurs des termes utilisés jusqu'à ce jour. Il s'agit de **kuon-* ou **kun-*, grec *kúōn*, irl. *cú*, got. *hun-d-s*, all. *Hund*, et, dans les langues *satem*, letton *sun-s*, arm. *šun*, skrt *cvan-*, avest. *span-*, médique *spa-ka*, d'où le russe *sobaka*.

Le chat, au contraire, n'a été régulièrement attaché à la maison que très tard, en Grèce et à Rome après la période classique. Le mot qui le désigne alors a l'allure d'un mot occidental pré-indo-européen avec, à l'initiale, [k] ou [g], une voyelle [a] et [t], au départ géminé entre voyelles, comme l'attestent l'ital. *gatto*, l'irl. *catt*, angl. *cat*, all. *Katze*, le vieux slave *kot-ŭka*, etc. Le chat n'étant pas là, les souris devaient danser, et, pour elles, on peut restituer une forme **mūs* qui est celle du grec *mûs*, du latin, du vieil-anglais et du vieil-islandais *mūs*, en russe régulièrement *mys'*.

Parmi les espèces sauvages, nous avons déjà rencontré le loup dont

la forme ancienne **wl̥k^{w}o-* alternant avec **luk^{w}o-* est attestée partout, sauf en celtique.

Pour l'ours, on reconstruit *H_2*erkto*, H_2*r̥kto* bien attesté ; en gr. *árktos*, skrt. *r̥kša-*, lat. *ursus*. Ce mot, qu'on rapproche de skrt *rakṣas-* « dommage », est peut-être, lui-même, le successeur d'un mot plus ancien, éliminé par interdiction superstitieuse de vocabulaire. Il a été remplacé en germanique par un mot qui veut dire « le brun » : angl. *bear*, all. *Bär*, et en slave par un composé du sens de « mangeur de miel », russe *medv-ed*. Le miel jouait certainement un grand rôle dans l'alimentation préhistorique et était à la base d'une boisson alcoolique, l'hydromel. Il y a, pour « miel », deux formes en compétition : **melit*, avec grec *méli*, lat. *mel*, irl. *mil*, got. *miliþ*, et **med^{h}u* qui désigne aussi l'hydromel ; avec le sens de miel, on a lit. *medus*, le russe *med* (*medv-* dans *medved* « ours » conservant la trace de l'ancien *-u*) ; avec le sens d'hydromel, on trouve irl. *mid*, angl. *mead*, dan. *mjød*, all. *Met*, lit. *midus* et russe *med*. On ne sait pas exactement à quoi se réfère la forme avestique *maδu-*, peut-être le vin, comme parfois le grec *méthu*. Le vin est certainement d'apparition tardive et sa désignation en Europe dérive partout du latin *uīnum*. Il faut mettre à part le grec où *krasi*, proprement « mélange », a remplacé l'ancien *(w)oînos* représentant, comme lat. *uīnum*, un emprunt à quelque langue méditerranéenne.

L'antiquité des mots désignant le loup et l'ours indique que dès l'époque la plus ancienne, les peuples indo-européens ont vécu dans des contrées où ces animaux étaient de fréquente occurrence. Mais on ne saurait tirer argument du fait que toutes les désignations du lion paraissent dériver de la forme grecque, pour affirmer que le lion était inconnu dans les régions où étaient fixées les populations de langue indo-européenne antérieurement à leur dispersion. Il pourrait avoir suffi, pour qu'il soit éliminé de la nomenclature zoologique, que les différents peuples, dans leur migration, aient dû séjourner dans des régions où le lion n'existait pas ou avait disparu. La légende parle de lions en Grèce, mais à constater l'importance du rôle qu'a joué cet animal dans l'imaginaire des peuples d'Europe depuis l'antiquité, on pourrait se demander si le Lion de Némée, tué par Héraclès, ne ferait pas partie des mythes. Toutefois, sa présence dans la péninsule balkanique est signalée par Aristote. Pour expliquer les formes prises par le mot à travers l'Europe, il faut partir, non de la forme grecque *léōn* (génitif *léont-os*), mais de la forme latine *leo*, *leōnis*. C'est fréquemment la forme du nominatif latin *leo* qui explique les diverses formes, ce qui fait supposer un emprunt livresque. C'est le *-o* de *leo* qui a donné le *-w-* de l'allemand *Löwe*, du néerlandais *leeuw* et, de proche en proche, le *v* du lit. *levas* et du russe *l'ev*.

Parmi les félidés, le lynx doit retenir notre attention. Son nom est attesté de l'Irlande à l'Arménie, sous une forme dérivée de **luk-*,

racine que l'on trouve dans gr. *leukós* « brillant », épithète souvent appliquée aux yeux du lynx : le moyen-irlandais présente *lug*, le suédois *lo* (< **luχa-*), l'all. *Luchs*, le lit. *lúszis*, le russe *rys'* avec un *r-* sur l'analogie de *rvat'* « déchirer ». Le grec a innové avec son *n* infixé dans *lúgx* [luŋks], gén. *lugk-ós* [luŋkos]. Le latin *lynx* est, bien entendu, emprunté au grec. La désignation populaire de l'animal est en français, *loup-cervier* (« qui s'attaque aux cerfs »), et l'on peut se demander si le nom du loup (**wl̥kʷo-*) n'a pas été influencé en grec (*lúkos*) et en latin (*lupus*) par celui du lynx. Le lynx, rare aujourd'hui, a dû être un hôte fréquent des forêts de l'Europe moyenne et septentrionale.

La flore.

De façon générale, la terminologie relative aux espèces animales ou végétales tendrait à faire croire que l'Europe moyenne telle que nous la connaissons, était l'habitat traditionnel des « Indo-Européens ». Mais il ne faut pas oublier que ce sont probablement les sources de notre documentation qui influencent notre jugement en la matière. On a vu, à propos du radical **bhāgo-*, que ce mot a pu désigner l'arbre qui, avant l'expansion néolithique, fournissait un élément important de l'alimentation humaine ou animale, sous la forme de la faine ou du gland. Dans un cas il s'agissait du hêtre, dans l'autre du chêne. Le nom du chêne en Gaule, **kassano-s*, qui a donné le mot français, a toute chance de représenter régulièrement un plus ancien **kastano-s*, d'où le latin *castanea* qui désigne la « châtaigne », ressource alimentaire qu'on a rapprochée du gland et de la faine. Le mot *faine* dérive de l'expression (*glans*) *fāgīna*, c'est-à-dire « gland de hêtre », *fāgīna* étant naturellement dérivé de *fāgus* « hêtre ». Quant au mot *gland*, il est attesté de l'italique à l'arménien, en passant par le balte, le slave et le grec par des formes dérivées d'un radical *gel°*H_2 ou *gle*H_2*-*. Pour illustrer la façon dont l'utilisation que l'homme fait des produits naturels détermine l'évolution du sens attribué à une forme, on rappellera que *corn, korn,* dans les langues germaniques, ne s'applique pas à une espèce déterminée, mais à celle qui est, dans chaque région, la céréale de base : aux Etats-Unis, *corn* désigne le maïs, en Irlande, l'avoine, en Angleterre, le froment ; en Allemagne, *Korn* s'applique au froment au sud, au seigle au nord.

L'agriculture.

L'agriculture s'est développée tardivement, parallèlement à l'élevage. L'ameublissement de la terre, première condition pour assurer la reproduction artificielle des espèces botaniques, a pu être effectué

initialement au moyen d'instruments du type de la bêche de bois d'un type analogue à celle qui est connue, en basque, sous le nom de *laia*. L'espagnol a *layar* pour « bêcher ». Il est possible que le verbe issu d'une racine H_2*ereH*$_3$ que nous traduisons par « labourer » et qui a en latin la forme *arāre* ait, au départ, désigné l'opération réalisée au moyen de cet outil primitif. Mais l'instrument dont le nom, lat. *arātru-m*, français *araire*, est dérivé de cette racine au moyen d'un suffixe qui peut varier d'une langue à une autre, est bien celui qui évoluera pour aboutir à ce qu'on désigne comme une charrue. Le fait que les mots de cette racine n'apparaissent pas en indo-iranien a été interprété comme un oubli dû à une longue période de migration. Mais il se pourrait aussi que l'invention de l'araire se soit faite en Europe postérieurement au départ des Aryens et que l'outil se soit diffusé, par exemple, à partir du bassin du Danube, depuis les Slaves, à l'est, jusqu'aux Irlandais, à l'ouest. A noter que le tokharien connaît *āre* pour la charrue.

Le mot *joug* est attesté partout et généralement comme un neutre en *-o*, par exemple, dans le hittite *yugan*, le skrt *yugám*, le grec *zugón*, le latin *iugum*, ce qui semble indiquer que l'attelage des bœufs remonte à une très haute antiquité. Il est difficile de préciser à quelle fin première, labourage ou traction du chariot dans les déplacements des nomades. Le terme n'est pas isolé. La racine **yeug-* dont il est tiré est largement représentée par des verbes du sens de « joindre ». Le fr. *joindre* lui-même, avec ses dérivés, comme *conjoint*, et la foule des formes « savantes » qui l'accompagnent : *jonction*, *conjonction*, *conjugaison*, *juxtaposer*, etc., illustrent bien le destin extraordinaire de cette racine.

Pour la notion de sillon, il existe une forme largement attestée, celle du latin *līra* qui remonte à un radical **loiseH*$_2$ > **loisā* représentée par le vieux-slave *lěxa* et le lit. *lyse* avec le sens de « plate-bande », le vieux-haut allemand *(wagan)leisa* « ornière », all. *Gleise* « ornière, voie (ferrée) », de *(Ge)leise*. *Délirer*, du latin *dēlīrāre*, c'est proprement « quitter le sillon », « dérailler », au sens figuré du terme. En germanique, avec un accent sur la finale qui a fait passer régulièrement *-s-* à *-z-*, puis à *-r-*, on aboutit à **lairō*, d'où l'all. *Lehre* « enseignement », l'angl. *lore*, comme dans *folklore* « étude des comportements populaires » ; le gotique, qui a souvent rétabli le *-s-* par analogie, présente *laisjan* « enseigner ». Pour « apprendre », on a un dérivé en *-na-* formé sur le vocalisme zéro **lis-*, devenu **liz-* puis **lir-*, dans l'all. *lernen* et l'angl. *learn*. Un autre dérivé est got. *laistjan* « suivre », c'est-à-dire « suivre la voie », représenté en allemand par *leisten* « créer, accomplir » et par l'anglais *last* « durer ». Le substantif correspondant est got. *laists* « trace du pied », d'où l'all. *Leisten* et l'anglais *last* qui désignent la forme à chaussure.

Pour la notion de semer, on trouve un peu partout la racine *seH_1 bien représentée en latin dans *sē-men* « semence » et dans *sē-u-i* « j'ai semé », avec un *-u-*, comme marque du parfait, analogique des radicaux en -H_3 ; cet élément est devenu partie constitutive du radical dans le vieil-anglais *sāwan* > anglais *sow* « semer ». Il n'y a guère que le grec pour faire bande à part avec la racine **sper-* représentée dans les deux mots du sens de semence, *spérma* et *spóros*, empruntés sous les formes de *sperme* et *spore*. Le français *semer* représente le latin *sēmināre*, dérivé de *sē-men*. Le rapprochement avec le hittite *sai-* suggère, comme sens primitif, « enfoncer ». Le terme se référerait donc aux semailles grain à grain et non à la volée. On a, de nouveau, affaire à une racine très prolifique.

Pour désigner le grain produit par les céréales, on trouve une base *yewo-*, attestée en indo-iranien sous la forme *yava-* en sanskrit et en avestique, qui finalement prend le sens d' « orge », et également en balte avec le pluriel lituanien *javai* et en grec avec *zeiaí*, pluriel également, désignant une variété de froment. La forme la plus répandue en Europe est celle qui se présente en latin comme *grānum* « grain » ; en germanique, elle est devenue *corn*, en anglais, *Korn*, *korn* en allemand et en scandinave, *zrŭno* en vieux-slave, *žirnas* en lituanien. Le mot, nous l'avons vu, désigne, en général, la céréale la plus répandue, et il peut même s'étendre, comme en balte, jusqu'aux pois. On trouve, parallèlement, des désignations spécifiques, pour l'orge notamment, avec *$g^hersd^{(h)}$- représenté par l'all. *Gerste*, en latin *hordeum*, le grec *krīthḗ* avec une voyelle inattendue, l'albanais *driθ*. On en rapproche l'arménien *gari*.

Les saisons.

Pour les saisons, on pourrait s'attendre à ce qu'on oppose simplement la belle à la mauvaise : même si nous opérons avec quatre saisons, nous connaissons, en matière vestimentaire, des demi-saisons qui correspondent au printemps et à l'automne. Toutefois, dans les langues anciennes, on trouve des formes communes pour l'hiver, le printemps et l'été. Les désignations anciennes de l'hiver remontent à une racine g^hei-, le plus souvent suivie de *-m-* : grec *kheimṓn*, lat. *hiems*, avec un adjectif *hībernus* qui suggère une ancienne prénasalisée ^mb avec un traitement final *-m* et un traitement interne *-b-*, irl. *gam* (avec un *-a-* analogique de celui de *sam* « été »), lit. *žiema*, vieux-slave et russe *zima*, skrt védique *himā-*. La même racine se retrouve, pour désigner la neige, dans le grec *khiṓn*, l'arménien *jiun* et le sanskrit *himā-*. Il n'y a guère que le germanique pour présenter une désignation divergente de la saison froide, avec got. *wintrus*, all. *Winter*, angl. *winter*, dan. *vinter*.

Pour le printemps, on a une base *wesn-*, avec *-r* au lieu de *-n* à la finale, représentée par grec *éar* (< **wear* < *wesr̥*), latin *uēr*, d'où l'ital. et l'esp. (*prima*)*vera*, irl. *err-ach* (qui a dû alterner, selon la position dans la phrase, avec **ferr-ach*), suédois *vår*, letton *vasara* qui a pris le sens d'été, vieux-slave et russe *vesna*.

Pour l'été, la forme de base est **se/om-* avec, dans certains cas *-a-* au lieu de *-e/o-*. Elle peut aussi valoir pour l'année tout entière, comme skrt *sama-* et l'arménien *am* ; l'avestique *hama* désigne l'été. Pour l'été, le vieil-irlandais a *sam*, le gallois *haf*, le breton *hañv*, toutes formes remontant à **samo-*. En germanique, un suffixe *-ro-* accentué s'est ajouté à la base **sem-*, d'où **s°mro-* et les formes contemporaines angl. *summer*, all. *Sommer*, suédois *sommar*. Mais l'été a été également désigné comme la saison chaude : lat. *aestas*, *aestāt-is*, d'où *été*, qui dérive d'une racine *H_2eid^h attestée dans grec *aithō* « brûler ».

La Fontaine emploie « au temps chaud » pour l'été. L'équivalent de *chaud-temps* est, en franco-provençal, la désignation régulière de cette saison. On rapprochera le patronyme *Chautemps*.

L'automne est à part en ce que chaque branche de la famille a, pour le désigner, une ou plusieurs formes particulières, comme l'all. *Herbst*, dan. *høst*, en face de l'anglais *harvest* qui, sous la pression de son concurrent *autumn*, imité du français, n'a gardé que la valeur primitive de récolte, celle qu'on retrouve dans le latin *carpo* « je cueille ». De la même racine, on a le grec *karpós* « fruit ». L'anglais d'Amérique a, pour « automne », le mot *Fall* « chute (des feuilles) » en face de *Spring* « printemps », qui reflète le jaillissement de la sève (*spring* « bondir, jaillir ») et la poussée de la végétation.

Les points cardinaux.

On dira ici quelques mots des points cardinaux. Pour le nord, on ne trouve pas de formes communes à plusieurs branches. Il est souvent désigné par le nom du vent froid qui en provient, comme *boréas* en grec. La forme *nord* est, comme les autres points cardinaux français, empruntée au vieil-anglais où l'on a *norð*, qu'on interprète comme voulant dire « à gauche » (osque et ombrien *nertro-* « la gauche »), le nord étant à gauche lorsqu'on regarde le soleil levant. Il n'y a guère plus de concordances, à date ancienne, en ce qui concerne le sud. On le trouve dans le sanskrit *dakšina* désigné comme « à droite », mais c'est aussi là où se trouve le soleil au milieu du jour, d'où, en français, *midi* (avec *-di* pour « jour » comme dans *lundi*, *mardi*...), en provençal et en franco-provençal *mi-jor(n)*. Ce peut être simplement le côté du soleil, comme en germanique, où on trouve pour « sud » un dérivé de la racine **sun* qui désigne le soleil : *sun-þ-* d'où le vieil-anglais *sūð*. Le

français a emprunté cette forme assez tôt pour que le *ū* [u :] y passe à [y]. Les formes *sud, sur, sud* des autres langues romanes ont été empruntées soit par écrit, soit assez tôt pour justifier leur [u]. L'allemand *Süd* est emprunté au néerlandais, la forme originelle *sund* étant attestée dans l'onomastique : la partie méridionale de l'Alsace est désignée comme le *Sundgau* « canton du sud ».

Pour l'est, on a une forme ancienne qui désigne aussi l'aurore et qu'on reconstruit comme **ausos*. On a, chez Homère, *ēṓs*, en lat. *aurora*, lit. *ausra* « aurore », en germanique *aust-*, d'où régulièrement, en vieil-anglais, *ēast* qui a été emprunté par le français sous la forme *est* et adopté en italien (*est*) et en espagnol (*este*). Le *aust* ancien se retrouve dans le français *Autriche*, en face de la forme allemande évoluée régulièrement en *Ost* (*öst-* dans *Österreich* « Autriche »). On notera, pour « Pâques », angl. *Easter*, en quelque sorte « l'aurore de l'année », et le pluriel allemand, de même sens, *Ostern*. Mais l'est est souvent désigné comme la direction où le soleil se lève : c'est la valeur du grec *anatolḗ*, d'où *Anatolie* qui est le nom que les Grecs donnent à l'Asie mineure. Le latin *oriens*, fr. *orient*, vaut exactement « levant ». En français, le *Levant* désigne les pays de l'est du bassin méditerranéen. Le russe *vostok* (connu par *Vladivostok* « maîtrise de l'Orient ») « est » suggère la montée du soleil.

L'ouest peut être « le soir », comme le grec *hespera* qui a les deux sens. C'est souvent la « chute du soleil », l'*Occident*, ou encore « là où il se place », le *ponant*. Les formes germaniques en *west-* sont à rapprocher du skrt *avas* « vers le bas » et du premier élément du grec *hespéra* (pour **wespera*) et du latin *uesper* « soir », d'où, en français, les *vêpres*. Le français *ouest*, qui reproduit la forme anglaise *west*, a été reproduit, oralement, en espagnol, sous la forme *oeste* et graphiquement en italien comme *ovest*, avec une méprise sur la valeur de *-u-* lorsqu'au XVIe ou au XVIIe siècle, on s'est avisé de distinguer entre *u* et *v*.

Les techniques domestiques.

A propos des techniques domestiques, il faut signaler que la couture en est une très ancienne, avec la racine **suH-*, **siuH-* présente un peu partout, sauf en celtique et en grec. Une racine **(s)neH₁-* qui a le sens de filer, puis de tisser et d'entrelacer, a fourni, au latin, *neo*, parfait *nē-u-i*, au grec, *néō*, futur *nē-s-o*. Le correspondant formel vieux-haut-allemand *nāan*, all. *nähen*, a le sens de coudre. L'aiguille est en gotique *nē-þ-la*, en anglais *needle*, en allemand *Nadel*. Il faut rapprocher cette base de **(s)new-* qui apparaît dans les désignations du tendon. Le tendon a dû jouer un rôle important dans la couture primitive. Cette racine est attestée dans le skrt. *snāvan-* « lien, tendon, cordon », dans

le tokharien (dialecte B) *ñaura* « nerfs », en grec dans *neûron* « fibre, corde, nerf » et *neurā́* « corde de l'arc » On la retrouve, avec des traitements phonétiques particuliers, dans le lat. *neruos* (> fr. *nerf*) et l'anglais *sinew* « tendon ».

Pour le tissage, on peut poser une racine H_2*ew-* avec des suffixes variables, en -b^h-, -d^h-, *-g-* : à partir d'une forme à vocalisme zéro *H_2ub^h-, on a le grec *(h)uph-aíno ;* *H_2web^h donne les formes germaniques, anglais *weave*, all. *weben*, d'où *Weber* « tisserand », patronyme très fréquent ; à partir de H_2wed_h- on trouve une désignation du vêtement : angl. *weed ;* les formes celtiques, irl. *figim*, gallois *gweu*, l'anglais *wick* « mèche », et, sans doute, le lat. *uēlu-m* « voile » (< *wesklo-*) supposent H_2*weg-*. Le latin *texere*, d'où dérive le français *tisser*, se dit aussi de la fabrication des cloisons par entrelacs et de la construction des navires, ce qui permet de rapprocher grec *tektōn*, skrt. *takšan* « charpentier », ainsi que des mots slaves, germaniques et celtiques désignant la hache.

Les métaux.

Même si des objets de métal ont été trouvés en abondance dans les sépultures qu'on a quelque raison de croire « indo-européennes », il n'est pas vraisemblable qu'une technique de fonte des métaux ait existé avant la dispersion. Le mot *fondre* qu'on fait remonter, au-delà du latin *fundere*, à une base g^h*eu-*, élargie en g^h*eu-d-*, veut dire très généralement « verser » sans spécification particulière. Sur la fabrication d'armes et d'outils à partir du silex, les formes indo-européennes ne nous renseignent guère. Il est intéressant de constater qu'au lat. *saxum* « pierre, rocher », correspondent des mots germaniques qui désignent des armes et des outils de bronze ou de fer. Les Saxons doivent leur nom à une arme dénommée *sax* (v. anglais *seax*, vieux-haut-all. *sahs*) qui était une dague. En danois, *saks* désigne les ciseaux. L'allemand *Messer* « couteau » remonte à *mezzi-sahs* > *mezzi-rahs*. Le premier élément y désigne la nourriture, comme le suédois *mat*, le vieil-anglais *mete*, l'anglais *meat* (noter *sweet-meats* « bonbons »), devenu aujourd'hui l'équivalent de « viande » sur le modèle de ce mot français qui, dérivé du latin *uīuenda*, s'appliquait à la nourriture en général.

Les objets de métal, obtenus par rapine ou par échange, ont dû, chez ceux qui nous intéressent, précéder la métallurgie, et il est fréquent que les noms des métaux ne remontent pas, d'une langue à une autre, à la même forme. Le cuivre, pur ou en alliage, pourrait être une exception : la forme représentée par lat. *aes* « bronze », qu'on retrouve, en français, dans le dérivé *airain*, est largement répandue de l'indo-iranien, skrt *áyas*, au germanique, got. *aiz*, v.-angl. *ār*, d'où

angl. *ore* « minerai », et laisse supposer un ancien **ayos-* (< *H_2*eyes-*). Lorsque, beaucoup plus tard, la distinction a été parfaitement établie entre le cuivre et ses alliages, le premier a reçu une désignation dérivée du nom de Chypre, source importante du métal, d'où fr. *cuivre*, angl. *copper*, all. *Kupfer*. Le grec *khalkós* (< *g^h*l̥*g^h*o-*) est à rapprocher des formes baltes, lit. *gelžis*, et slaves, russe *železo* (< *g^h*elg*^{h}-) qui désignent le fer, où l'évolution *satem* n'a pas joué pour la première consonne (le *ž* russe résulte d'un palatalisation secondaire). Il s'agit certainement d'un emprunt assez ancien à une langue non identifiée.

Pour les autres métaux, on retrouve des formes analogues dans des zones géographiquement limitées. Pour l'argent, par exemple, on rapprochera les formes germaniques, angl. *silver*, all. *Silber*, dan. *sølv*, de formes slaves comme russe *serebro*, ou baltes, comme lit. *sidabras*. Ceci vaut également pour l'or où les langues du nord présentent des formes dérivées d'une base *g^h*l̥tó-*, *g^h*élto-*, *g^h*olto-*, d'où got. *gulþ*, angl. *gold*, all. *Gold*, dan. *guld* et, avec traitement *satem*, letton *zèlts*, v. slave *zlato*, russe *zoloto*. On en rapproche toutefois des formes indo-iraniennes de même sens, comme skrt *hiraṇyam*, avec référence à la couleur jaune.

Ailleurs, l'argent a été désigné comme le métal brillant, d'une racine *H_2*erg-*, d'où lat. *argentum*, irl. *argat*, arménien *arcat'*, tokharien A *ārkyant*, skrt *rajatám*, grec *árguros*.

Pour l'or, on a une base *H_2*euso-* représentée par lat. *aurum*, d'où non seulement les formes romanes, comme fr. *or*, mais également les équivalents celtiques, irl. *ór*, gall. *awr*, et, d'autre part, le vieux-prussien *ausis* et le lituanien *auksas*. On a, en tokharien A, pour « or », *wäs* (< *H_2*wes-*) à côté de *wsi* « jaune ».

En ce qui concerne le fer, son apparition tardive est reflétée par la variété des termes qui le désignent. Du lat. *ferrum*, on rapproche l'angl. *brass* « laiton », et l'on peut restituer une base commune *b^h*e/ors-*; mais on rencontre des formes évidemment apparentées dans des langues sémitiques. L'utilisation d'une même forme pour des métaux différents a été illustrée ci-dessus par l'emploi, en slave et en baltique, de la forme traditionnelle pour le cuivre en référence au fer. Dans les langues germaniques, les désignations du fer sont empruntées au celtique, mais à deux étapes différentes de l'évolution de la forme **īsarno-* attestée dans l'onomastique gauloise. On trouve la forme avec *-s-* dans l'adjectif gotique *eisarn* et l'all. *Eisen* ou, mieux, dans l'adjectif correspondant *eisern*. En scandinave, on a également, à date ancienne, *īsarn*, mais très vite *jarn*, dan. *jern*, suéd. *järn* qui reproduit une forme celtique plus récente, irl. *iarn*, avec chute régulière de *-s-* entre voyelles, donc **īarno*. L'anglais *iron*, ['aiən] en britannique, ['aiərn] en Amérique, remonte également à **īarno-*.

Pour le plomb, les formes lat. *plumbum*, grecques *mólubdos*, *mólibos*,

bólimos, *bólibos*, germaniques, angl. *lead*, all. *Lot* (**loud*^{h}*o-*) pourraient être des emprunts à une même base non indo-européenne, quelque chose comme *^{m}*blu*^{m}*b-do-*, forme expressive, donc très instable, suggérant la grande densité du métal.

Comme on le voit par ce qui précède, les désignations des métaux ont voyagé avec les métaux eux-mêmes et, du fait des circonstances particulières des emprunts, les correspondances phoniques sont le plus souvent approximatives.

Le négoce.

Lorsqu'on cherche, dans les langues qui nous concernent ici, les traces de préoccupations commerciales, on les rencontre à basse date, dans des sociétés complexes où est répandue la division du travail, et l'on relève leur diffusion par emprunt dans des langues qui, par contacts, s'initient aux échanges en tant qu'occupation professionnelle. En latin, le mot qui veut dire « acheter », *emo*, avait précédemment le sens de « prendre » ; il existait traditionnellement un terme pour désigner l'achat, *uēnum*, et la valeur « vendre » a été obtenue par combinaison de ce terme avec le verbe *dō* « donner », sous la forme *uēnum dō* « je donne à achat », passée à *uendo*. Le mot *negōtium*, qui désigne le commerce, est formé du préfixe négatif *nec-* et d'*-ōtium* qui désigne l'oisiveté. Le commerce est, en quelque sorte, une façon de remplir son temps, de s'activer un peu. A partir de *busy* « occupé », l'anglais a dérivé *business* « affaire ». En français, *affaires* vient naturellement de *à faire*, au départ, sans doute, dans *avoir affaire*. La forme latine *negōtium* est un calque du grec *a-skholía* « manque de loisir », puis « affaire », du préfixe négatif *a-* et de *skholḗ* « loisir », d'où le latin *schola* et notre *école*. Les premiers commerçants qu'ont connu les Germains ont été les cabaretiers d'origine romaine, les *caupōnes*. C'est de *caupō* que dérivent le got. *kaupon* « faire commerce » et les mots allemands *kaufen* « acheter » et *verkaufen* « vendre ». Le slave (russe *kupit'* « acheter ») a reçu la forme du germanique. Le danois a *købe* pour « acheter », mais comme l'anglais, une forme plus traditionnelle pour « vendre » : dan. *sælge*, angl. *sell*, dont le sens premier est celui de l'équivalent gotique *saljan* « faire un sacrifice », c'est-à-dire se séparer d'un objet de valeur. L'anglais *buy* « acheter », qu'on rapproche du got. *bugian*, de même sens, pose des problèmes de forme : la forme ancienne est *bycgan*, d'où on attendrait **bidge* ; son prétérit *bought* correspond à l'élément *-bauhts* du gotique *anda-bauhts* « rédemption » et au skrt. *bhúkti-ṣ́* « libération ». Ceci tend à faire croire que le mot s'appliquait initialement à la libération des prisonniers ou des esclaves par rachat. La notion d'une aide apportée par l'achat est encore sensible dans certains emplois

modernes de *buy*, comme *buy me a drink!* « paie-moi un verre ! ». Dans une société où, dans chaque cellule, on était censé fabriquer tous les outils nécessaires à la subsistance et où des esclaves étaient disponibles à cette fin, c'était des bras et des compétences techniques qu'on achetait plutôt que des objets. Ce qui, dans ce contexte, valait pour les hommes, valait aussi pour le bétail. En grec, *andrápoda* « esclaves », un neutre, littéralement « pieds d'homme », paraît calqué sur *tétrápoda* (accusatif) « quadrupède ». Le rachat, sinon l'achat, correspondait souvent à une libération, celle qu'implique le terme *rédemption* (*red-* équivalent de *re-* + *emo-* « acheter »).

Les couleurs.

La vision qu'on a du monde, au sens le plus matériel de ce terme, est conditionnée par la perception des couleurs. Contrairement à ce que se figurent la plupart des gens, les couleurs distinctes ne sont pas une donnée naturelle. L'arc-en-ciel présente un continu qu'on analyse en fonction d'une gamme chromatique qui est fournie par la langue qu'on utilise. Sans aller très loin, on peut signaler que les Bretons et les Gallois n'ont qu'un seul mot, *glas*, qui recouvre les zones françaises du bleu et du vert, sans préjudice du gris. Dans bien des langues, que nous sommes tentés de qualifier de « primitives », on ne distingue, du noir et du blanc, que le seul rouge, et telle semble bien avoir été la situation en indo-européen commun. D'un bout à l'autre du domaine, on rencontre pour désigner cette couleur, des formes dérivées d'une racine **reud*^{h}, latin dialectal *rūfu-s*, irl. *rúad*, breton *ruz*, got. *rauþ-s*, angl. *red*, all. *rot*, fréquemment avec un suffixe *-ro-* connu dans le skrt. *rudhirá-s* et le grec *(e)ruth-ró-s*, le latin *rube-r*, le vieux-slave *rŭd-rŭ*. D'autres désignations du rouge ou de ses variétés dérivent de techniques de teinture qui faisaient intervenir diverses bestioles, comme certains vers, d'où par exemple, le français *vermeil*, du latin *vermiculu-m* « petit ver » ou le vieux-slave *črĭvenŭ*, de *črŭvĭ* « ver ». La base *reud*^{h}- se retrouve en latin, en germanique, en balte et en slave pour désigner la rouille, et il est vraisemblable que *rouge* a longtemps évoqué plutôt la couleur de la brique que celle du coquelicot, d'où le besoin de termes comme *vermillon* et *écarlate* pour préciser qu'il s'agissait d'un « rouge vif ».

Pour le blanc et le noir, qui nous paraissent aujourd'hui, en Occident, deux perceptions bien caractérisées, avec un terme unique pour chacune, on trouve anciennement, en latin, par exemple, des formes différentes pour l'un et pour l'autre selon qu'il s'agit d'un blanc ou d'un noir terne ou éclatant : en latin on a, d'une part, *albus* pour la blancheur de l'*aube*, *ater* pour la noirceur du charbon ; d'autre part, *candidu-s* pour la couleur de la toge, frottée à la craie, des

« candidats » aux fonctions publiques et *niger* pour le noir en général. Ces deux derniers vont ensemble : *nigrum in candida uertere* « changer le noir en blanc », c'est-à-dire « tromper son monde ». L'examen des désignations d'autres couleurs montre que ce qui frappait les anciens était essentiellement la luminosité et non la couleur proprement dite. En grec, *leukós*, tout comme *glaukós*, veulent dire « brillant » avant de se spécialiser, le premier, dans la blancheur, le second, avec une valeur couvrant à peu près le domaine du celtique *glas*, donc du bleu au vert et au gris. On rattache à une même base *b^hleH_3- aussi bien les mots germaniques pour « bleu » et le français *bleu* lui-même, que les mots latins *flāuos* « jaune » et *flōrus* « blond ». La même racine, avec un vocalisme différent, *b^helH_3 pourrait être à la source des mots baltes et slaves pour « blanc » : lit. *baltas*, vieux-slave *bělŭ*.

La plupart des noms de couleur dérivent de ceux d'objets caractérisés par la couleur en question. Le vert, par exemple, est la couleur de la végétation et l'anglais *green*, l'all. *grün* ne sauraient être séparés de *grow* « croître », non plus que lat. *uiridis*, d'où fr. *vert*, de la racine attestée dans le lituanien *veisti* « propager ». Un cas intéressant est celui de la prune et de son antécédant sauvage, la prunelle. La base est **sle/oiw-*, alternant avec **sle/oik-*; les langues slaves ont *sliva* pour la prune, d'où l'alcool de prunes connu sous le nom de slivovitsa ; en germanique, on a, en vieux-haut-allemand *slēha*, *slēwa*, d'où l'allemand *Schlehe* « prunelle ». L'équivalent anglais est *sloe*. En latin, avec chute régulière du *s*- initial, on a *līu-ēo* « être blême » et l'adjectif *līuidus* « livide ». Le celtique présente l'irl. *li* et le gallois *lliw* qui veulent dire « couleur ».

La notion même de couleur paraît d'origine récente. Les mots qui assument aujourd'hui cette valeur remontent, le plus souvent, à des formes qui désignaient l'apparence, la surface, la couverture, le plumage, le pelage. En latin, par exemple, *color*, est à rapprocher du *-cul-* de *occulere*, d'où *occulter*, et du *cēl-* de *cēlāre* « cacher ».

On voudrait pouvoir poursuivre cette revue de ceux des champs sémantiques où la comparaison des langues autorise à restituer des formes plus anciennes et à imaginer à quoi elles se référaient. Mais on arriverait vite à des rapprochements qui feraient aux hypothèses une place encore plus large que ci-dessus, et ceci surtout si l'on désire, comme nous le faisons ici, accéder aux valeurs anciennes, au-delà donc des significations que suggèrent les textes avec lesquels on opère. Il faut surtout ne jamais oublier que les langues qui dérivent d'un même idiome plus ancien vont, pendant des siècles, voire des millénaires évoluer parallèlement, jusqu'à ce que se fassent sentir les effets d'influences particulières auxquelles chacune d'entre elles aura été exposée.

Ce parallélisme évolutif tient, tout d'abord, à ce que la structure de toute langue porte, en elle-même, les germes de son évolution. De façon plus explicite, disons que lorsqu'il s'agit, pour les locuteurs, de satisfaire leurs nouveaux besoins de communication, ils tentent de le faire avec les ressources que leur offre la langue qu'ils ont reçue de leurs aînés. Ces ressources ne consistent pas en un conglomérat hétéroclite de vocables différents, mais en un complexe d'habitudes articulatoires, syntaxiques, dérivationnelles et autres qui peut résister avec succès aux ingérences extérieures. Sans doute, ce complexe n'empêchera-t-il pas l'évolution, mais il l'orientera constamment. Un jour viendra, sans doute, où la somme des adaptations aux nouvelles conditions morales, sociales et économiques sera telle que la langue, dans son aspect et dans sa structure, sera devenue méconnaissable. Seul un examen attentif pourra alors retrouver les traces d'identités originelles.

A côté des résistances de la structure linguistique, il y a, pour expliquer le parallélisme évolutif, l'existence de tendances générales qui conditionnent le devenir de l'humanité. Ce n'est pas un hasard si, au cours des quelque douze cents ans qui se centrent autour du début de notre ère, sont apparues ce qu'on désigne comme les religions de salut, du bouddhisme et du mazdéisme au christianisme et à l'islam, celles qui ne sont plus, comme le paganisme, la simple constatation de ce qui semble régir l'évolution du monde et le destin de l'homme, mais qui établissent l'opposition entre le bien et le mal. Du latin aux langues romanes qui en sont issues, il n'y a pas seulement la palatalisation des dorsales, l'effritement, puis l'élimination des déclinaisons, le renouvellement du vocabulaire, mais, au-delà de la conservation des termes anciens, les valeurs nouvelles que leur a conférées la nouvelle foi, et ceci sur un modèle quasi identique dans toutes les parties de la Romania soumises à l'Eglise de Rome.

Si donc la comparaison des langues romanes ne saurait garantir une restitution des valeurs sémantiques exactes de la langue parlée à Rome avant notre ère, comment peut-on supposer que la comparaison du grec et de l'indien, parlés et écrits au premier millénaire avant Jésus-Christ, pourrait nous renseigner exactement sur le monde tel qu'il était perçu et vécu par un peuple des steppes, quatre ou cinq mille ans plus tôt ? Il nous faut en prendre notre parti, admettre les limitations que cela nous impose et nous bien garder de porter, sur la culture et les croyances que suggèrent nos reconstructions, des jugements de valeur qui mettraient, entre elles et nous, le voile de nos convictions et de nos préjugés.

TABLE DES CONVENTIONS GRAPHIQUES

Figurent en italique :

1°) les mots cités sous leur forme orthographique,
2°) les mots cités dans une translittération,
3°) les formes reconstruites précédées d'un astérisque.

Figurent en romain :

1°) entre barres obliques (/.../), les notations phonologiques (éléments distinctifs),
2°) entre crochets carrés ([...]), les notations phonétiques (réalité physique),
3°) entre guillemets, la traduction des formes citées.

Dans toutes les graphies représentées ici, en italique ou en romain, les lettres suivantes ont la valeur phonique qu'elles ont généralement en français : a, b, d, f, g (toujours dur), i, k, l, m, n, p, s, t, v, w (comme dans *watt*), z.

Cas spéciaux :

a et ɑ : si l'on doit distinguer entre les voyelles de *patte* et de *pâte*, [a] désigne la première et [ɑ] la seconde.

å indique une voyelle intermédiaire entre *a* et *o*.

æ indique une voyelle intermédiaire entre *a* et *è*.

c est souvent comme *ts* dans *tsar ;* č, comme *tch* dans *tchèque ;* ç comme *ch* dans l'allemand *ich*, version assourdie de *y* dans *yole*.

ch souvent équivalent de x ou de χ (cf. ci-dessous).

e comme *é* dans *pré* ou *è* dans *près*, mais si l'on spécifie ce dernier, on note [ɛ].

ə comme *a* dans l'anglais *villa* ou *e* dans l'allemand *Sage*.

h comme dans l'anglais *hand* ou l'allemand *Hand*.

ʰ, au-dessus de la ligne, indique que la consonne qui précède se prononce à

glotte ouverte comme *p* ou *t* en anglais, mais l'aspiration qui en résulte est plus perceptible qu'en anglais.

ḫ en hittite : cf. x ci-dessous.

H, dans les reconstructions, indique une consonne, dont la valeur précise est inconnue, qui disparaît en allongeant la voyelle qui précède.

ⁱ au-dessus de la ligne, indique le timbre *i* de la consonne qui précède. On trouve aussi, avec valeur analogue, ʸ ou ʲ.

j comme *dj* dans *adjoint* en anglais et en sanskrit ; en espagnol comme x (cf. ci-dessous) ; ailleurs, comme le *y* de *yole* ou marque de la mouillure de la consonne qui précède.

m, n, en général, ne nasalisent pas la consonne qui précède ; une voyelle nasale est marquée par le tilde : *õ* comme le *on* de *pont*.

o comme *o* dans *sot* ou dans *sotte*, mais si l'on spécifie ce dernier on note [ɔ].

ø comme *eu* dans *peu*.

[ɔ̃] est la voyelle qu'on entend souvent dans la première syllabe de *joli*.

r indique proprement un *r* roulé de la pointe de la langue, R un *r* roulé de la luette, ʁ (*r* renversé) le *r* ordinaire à Paris.

š comme *ch* dans *chante*.

u comme *ou* dans *pour*, sauf en grec et en néerlandais où il vaut *u* de *pur* et en suédois où il a une valeur analogue ; *ü* allemand comme *u* dans *pur*.

ᵘ au-dessus de la ligne indique le timbre *ou* de la consonne qui précède ; on trouve aussi ʷ avec valeur analogue.

ʷ au-dessus de la ligne indique que la consonne qui précède s'articule en même temps qu'un *w* comme dans *watt*.

x ou le grec χ, dans les translittérations et les transcriptions, comme le *ch* de l'allemand *ach* ou le *j* de l'espagnol *jota ;* les Français entendent un *r* sourd.

y comme dans *yole* dans les reconstructions ; en russe, un *i* avec langue rétractée ; dans les transcriptions, le *u* de *pur*.

ž comme *j* dans *jambe*.

θ ou þ comme le *th* de l'anglais *thin* ou le *c* de l'espagnol *cinco*.

ð comme le *th* de l'anglais *that*.

δ comme le *d* de l'espagnol *cada*.

β comme le *b* de l'espagnol *cabeza*.

γ comme le *g* de l'espagnol *hogar*.

ф est un *f* prononcé avec les deux lèvres.

ŋ comme *-ng* dans l'anglais *meeting*.

° placé au-dessus de la ligne dans les reconstructions indique une voyelle réduite de timbre imprécis.

̥ placé au-dessous d'une consonne indique qu'elle forme le sommet d'une syllabe.

̣ placé au-dessous d'une consonne indique, en sanskrit, que la consonne se prononce avec la pointe de la langue recourbée vers l'arrière.

ʼ placé après une consonne indique que cette consonne est palatalisée

(articulée en même temps qu'un *i*) ou palatale, c'est-à-dire articulée avec le dos de la langue contre le palais dur ; pour ce dernier type, on trouve aussi les notations [ṯ], [ḏ].

ˀ, fermeture glottale, s'entend au début de l'allemand *acht*.

ʾ, au-dessus de la ligne, indique que la consonne précédente se produit avec la glotte fermée.

ʕ, contraction pharyngale accompagnée de vibrations glottales.

ḥ, la même sans vibrations glottales.

˘, au-dessus d'une voyelle, indique qu'elle est brève.

¯, au-dessus d'une voyelle, indique qu'elle est longue.

:, après une voyelle, dans les transcriptions, indique qu'elle est longue.

Dans la translittération du grec, le iota souscrit est adscrit.

INDEX DES LANGUES ET DES PEUPLES

INDEX SÉLECTIF DES NOTIONS

INDEX DES AUTEURS

TABLE DES MATIÈRES

Ouvrage réalisé en photocomposition
par l'imprimerie BUSSIÈRE
et imprimé sur presse CAMERON
dans les ateliers de la S.E.P.C.
à Saint-Amand-Montrond (Cher),
le 3 septembre 1986

— N° d'impression : 3361-1944. —
Dépôt légal : septembre 1986.

Imprimé en France

LANGAGES ET SOCIÉTÉS
collection dirigée par Louis-Jean Calvet

La conquête du monde par les peuples de l'Occident a été longtemps ressentie comme étant dans la nature des choses. C'est au moment où elle rencontre des remises en question et des résistances efficaces que l'on commence à prendre conscience de la particularité du phénomène. En dépit de péripéties diverses, de conflits internes qui culminent aujourd'hui avec l'opposition des deux blocs, il s'agit bien d'une même expansion qui se poursuit depuis quelque six mille années.

Des peuples qui parlaient, au départ, une seule et même langue ont mis leur supériorité technique au service de la violence pour subjuguer leurs voisins de proche en proche. Cette langue commune, disparue, qu'on désigne comme l'indo-européen, on s'efforce de la restituer par la comparaison des langues connues, étayée par des hypothèses fondées sur la connaissance que nous avons aujourd'hui des idiomes les plus divers. En s'inspirant de récentes découvertes archéologiques, l'auteur a tenté d'esquisser la préhistoire de cette expansion et l'évolution de la langue des conquérants, depuis les formes que l'on reconstruit jusqu'aux formes attestées depuis la fin du deuxième millénaire avant notre ère. Il s'est efforcé de dépasser l'image statique que l'on donne en général de l'indo-européen commun en restituant la dynamique de son évolution.

L'Auteur : Professeur (émérite) Université René Descartes, directeur d'études Ecole pratique des hautes études. A publié de nombreux ouvrages de linguistique.

PAYOT, 106, boulevard Saint-Germain, PARIS

9 782228 140805
ISBN 2-228-14080-5

Michel Méline graphiste

150 F